

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: November 9, 2005, 12:50:22 ; Search time 214 Seconds
(without alignments)
2561.461 Million cell updates/sec

Title: US-10-724-972A-6352
Perfect score: 1690
Sequence: 1 GVESVRLKILSVIGLLFVL.....EEMAKELVLSKDKSKDK 335

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 81818359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US10724972/runat_03112005_141901_29159/app_query.fasta_1.519
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10724972@cgn2_1_69@runat_03112005_141901_29159 -NCFU=6 -ICPU=3
-NO_WMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	475	28.1	668	4	US-08-956-171E-355
2	475	28.1	668	4	US-08-781-986A-355
3	405	24.0	999	4	US-09-830-217-15
4	405	24.0	999	4	US-10-278-946-15
5	405	24.0	3775	4	US-08-956-171E-238
6	405	24.0	3775	4	US-08-781-986A-238
7	334	19.8	2115	4	US-08-956-171E-604
8	334	19.8	2115	4	US-08-781-986A-604
9	333	19.7	242	4	US-08-956-171E-2556
10	333	19.7	242	4	US-08-781-986A-2556
11	307.5	18.2	1059	4	US-09-107-532A-862
12	296.5	17.5	1056	3	US-09-134-001C-2655

C	13	296.5	17.5	3153	4	US-09-710-279-3348	Sequence 3348, Ap
	14	296.5	17.5	3267	4	US-09-710-279-4042	Sequence 4042, Ap
	15	296.5	17.5	3618	4	US-09-710-279-3564	Sequence 3564, Ap
	16	296	17.5	1056	4	US-09-543-681A-2690	Sequence 2690, Ap
	17	292	17.3	1044	4	US-09-710-279-461	Sequence 461, App
	18	292	17.3	1044	4	US-09-710-279-1267	Sequence 1267, App
	19	291	17.2	702	4	US-09-902-540-2208	Sequence 2208, Ap
	20	291	17.2	709	4	US-09-902-540-1583	Sequence 1583, Ap
	21	283	16.7	975	3	US-09-071-035-37	Sequence 37, Appl
	22	283	16.7	984	4	US-09-134-000C-937	Sequence 937, App
	23	278.5	16.5	978	4	US-09-543-681A-1393	Sequence 1393, Ap
	24	276.5	16.4	954	4	US-09-891-641-16	Sequence 16, Appl
	25	268	15.9	889	3	US-09-071-035-39	Sequence 39, Appl
	26	266	15.7	1041	4	US-09-602-787A-571	Sequence 571, App
	27	265	15.7	183	3	US-09-134-001C-2023	Sequence 2023, Ap
C	28	257	15.2	28194	4	US-09-902-540-1250	Sequence 1250, Ap
	29	254	15.0	1035	4	US-09-583-110-1062	Sequence 1062, Ap
C	30	253	15.0	9707	3	US-08-961-527-164	Sequence 164, App
	31	250	14.8	966	4	US-09-107-433-2206	Sequence 2206, Ap
	32	247.5	14.6	1119	4	US-09-602-787A-597	Sequence 597, App
	33	241	14.3	975	4	US-09-328-352-318	Sequence 318, App
	34	240	14.2	954	3	US-09-071-035-69	Sequence 69, Appl
	35	238	14.1	837	4	US-09-902-540-5089	Sequence 5089, Ap
	36	236	14.0	895	3	US-08-961-083-23	Sequence 23, Appl
	37	236	14.0	895	4	US-09-536-784-23	Sequence 23, Appl
	38	236	14.0	906	4	US-09-134-000C-2609	Sequence 2609, Ap
	39	228	13.5	868	3	US-09-071-035-71	Sequence 71, Appl
	40	218	12.9	966	4	US-09-302-6268-11	Sequence 11, Appl
	41	217	12.8	966	4	US-09-302-6268-9	Sequence 9, Appl
	42	217	12.8	888	4	US-09-107-433-1916	Sequence 1916, Ap
	43	210	12.4	900	4	US-09-710-279-1887	Sequence 1887, Ap
	44	205.5	12.2	930	3	US-09-134-001C-2830	Sequence 2830, Ap
	45	205.5	12.2	930	3	US-09-134-001C-2830	Sequence 2830, Ap

ALIGNMENTS

RESULT 1
US-08-956-171E-355
; Sequence 355, Application US/08956171E
; Patent No. 6593114
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
Gil H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman

BEST AVAILABLE COPY

REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 355:
SEQUENCE CHARACTERISTICS:
LENGTH: 668 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 355:
US-08-956-171E-355

Alignment Scores:
Pred. No.: 3,76e-45 Length: 668
Score: 475.00 Matches: 91
Percent Similarity: 86.89% Conservatives: 15
Best Local Similarity: 74.59% Mismatches: 16
Query Match: 28.11% Indels: 0
Gaps: 4

US-10-724-972A-6352 (1-335) x US-08-956-171E-355 (1-668)

Qy 210 LeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLys 229
Db 2 TTATTAGCACATCAAACTATTATATGTTGGACAAATTTTAAACGAACCTAGGATTTAAA 61

Qy 230 GluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeu 249
Db 62 AATGATTAAAGTACGATGTAACAAAGGTTTAAAGTAAATTTTAAAGGACCTTACTTA 121

Qy 250 GlnMetAnthrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn 269
Db 122 CAATTAGACACTGAACATTTAGCTGATTTAAATCCAGACGCTATGATCATTTATGACAGAT 181

Qy 270 LysAlaSerSerAsnGluProSerLeuLysGluLeuGluLysAspProValTyrLysLys 289
Db 182 CATGCTAAAAAGATTTCTGCTGAATTCAGAAAGTTTACAAAGATGCAACATGGAAAAAG 241

Qy 290 LeuAnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTyrAlaArgSer 309
Db 242 TTGAATGCAGTTAAAAAATAATCGCGTGATTTGTTGACCGTGATGTTGGCAAGATCT 301

Qy 310 ArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAsp 329
Db 302 CGTGCTTAATTTCTTCTGAAGAAATGCTAAAGAACTTTGTTGAATTTATCAAAAAAGAA 361

Qy 330 SerLys 331
Db 362 CAAAAG 367

RESULT 2
US-08-781-986A-355
Sequence 355, Application US/08781986A
Patent No. 6737248
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5355
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:

REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 355:
SEQUENCE CHARACTERISTICS:
LENGTH: 668 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 355:
US-08-956-171E-355

Alignment Scores:
Pred. No.: 3,76e-45 Length: 668
Score: 475.00 Matches: 91
Percent Similarity: 86.89% Conservatives: 15
Best Local Similarity: 74.59% Mismatches: 16
Query Match: 28.11% Indels: 0
Gaps: 4

US-10-724-972A-6352 (1-335) x US-08-956-171E-355 (1-668)

Qy 210 LeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLys 229
Db 2 TTATTAGCACATCAAACTATTATATGTTGGACAAATTTTAAACGAACCTAGGATTTAAA 61

Qy 230 GluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeu 249
Db 62 AATGATTAAAGTACGATGTAACAAAGGTTTAAAGTAAATTTTAAAGGACCTTACTTA 121

Qy 250 GlnMetAnthrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn 269
Db 122 CAATTAGACACTGAACATTTAGCTGATTTAAATCCAGACGCTATGATCATTTATGACAGAT 181

Qy 270 LysAlaSerSerAsnGluProSerLeuLysGluLeuGluLysAspProValTyrLysLys 289
Db 182 CATGCTAAAAAGATTTCTGCTGAATTCAGAAAGTTTACAAAGATGCAACATGGAAAAAG 241

Qy 290 LeuAnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTyrAlaArgSer 309
Db 242 TTGAATGCAGTTAAAAAATAATCGCGTGATTTGTTGACCGTGATGTTGGCAAGATCT 301

Qy 310 ArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAsp 329
Db 302 CGTGCTTAATTTCTTCTGAAGAAATGCTAAAGAACTTTGTTGAATTTATCAAAAAAGAA 361

Qy 330 SerLys 331
Db 362 CAAAAG 367

RESULT 2
US-08-781-986A-355
Sequence 355, Application US/08781986A
Patent No. 6737248
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5355
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:

```

: PRIOR APPLICATION NUMBER: 60/084,674
:
: PRIOR FILING DATE: 1998-05-07
:
: NUMBER OF SEQ ID NOS: 22
:
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 15
:
: LENGTH: 999
:
: TYPE: DNA
:
: ORGANISM: Staphylococcus aureus
US-09-830-217-15

```

Alignment Scores:	7,578-37	Length:	999
Pred. No.:	405.00	Matches:	113
Score:	51.74†	Conservative:	65
Percent Similarity:	32.88†	Mismatches:	132
Best Local Similarity:	23.96†	Indels:	13
Query Match:		Gaps:	12
DB:	4		

US-10-724-972A-6352 (1-335) x US-09-830-217-15 (1-999)

Qy	.8	Leu	Lys	Ile	Leu	Ser	Val	Ile	Gly	Leu	Leu	Phe	Val	Leu	Ile	Ala	Thr	Ala	A	Cys	Gly	27
Db	19	AT	TAA	AA	TGCTT---	GT	TG	TAC	GCTT	GCTT	TTC	TACT	TGGT---	TT	AGC	AGG	ATG	TAGT	72			
Qy	28	Asn	Asn	Ser	Ser	Ser	Asn	Ser	Ser	Glu	Ser	Ser	Lys	Asp	Gly	Val	Glu	Ile	Lys	His	47	
Db	73	GGG	AA	ATT	CAAA	TAA	CAAA	TAT	CAT	CTG	TATA	CAACA	AGA	TAA	GCGA	AA	CACTT	CAATT	TAA	CAAT	132	
Qy	48	Glu	Glu	Gly	Thr	Thr	Lys	Val	Pro	Lys	His	Pro	Lys	Arg	Val	Val	Val	Leu	Glu	Tyr	Ser	67
Db	133	GCA	AT	GGG	TAC	AACT	TGA	AACTTA	AAAG	GGAA	CAAC	CAAA	AGCG	TGT	TGTT	TAC	GCTAT	TAT	CAAGT	192		
Qy	68	Phe	Val	Asp	Ala	Leu	Val	Ala	Leu	Asp	Val	Lys	Pro	Val	Gly	Ile	Ala	Asn	Asp	Asn	Lys	87
Db	193	GCC	ACT	GAC	GC	TGCT	TAT	CTTT	AGTG	TTA	AACT	GT	TAG	GGT	GCT	GT	GTA	GAAT	CAT	CGG	ACA	252
Qy	88	Lys	Asn	Arg	Ile	Ile	Lys	Pro	Leu	Asp	Lys	Ile	Gly	Lys	Tyr	Thr	Ser	Val	Gly	Thr	107	
Db	253	CAAAA	CCG	CAA	ATT	TCG	AA	TAC	ATA	AAAA	TGATTT	AAAG	GAT	ACT	TAA	GAT	ACT	TAA	GAT	TG	TGCTCA	312
Qy	108	Arg	Lys	Gln	Pro	Asn	Leu	Glu	Glu	Ile	Ser	Lys	Leu	Lys	Pro	Asp	Leu	Ile	Ile	Ala	Asp	127
Db	313	GAAC	CT	TGC	ACC	TAA	CTTAG	AGG	AAAT	CTCT	TAAT	TAA	AA	CCG	GACTT	TAAT	TG	TCG	CGTCA	372		
Qy	128	Asn	Asn	Arg	His	Lys	Gly	Ile	Tyr	Lys	Asp	Leu	Asn	Lys	Ile	Ala	Pro	Thr	Ile	Glu	Leu	147
Db	373	AAAG	TT	AGAA	TG	AAAA	AGTTT	TAC	GTAT	CTA	AAAT	TAT	CTA	AAAT	TG	CA	CAAC	CA	CAAG	T---	426	
Qy	148	Lys	Ser	Phe	Asp	Gly	Asp	Tyr	Asn	Glu	Asn	Ile	Asp	Ala	Phe	Lys	Thr	Ile	Ser	Lys	Ala	167
Db	427	---	TCT	ACT	GAT	CAT	AGTTT	TT	CAAA	NTT	CAAA	---	GATA	CAAC	TAA	TAG	TTA	TG	GGG	AAAGCT	480	
Qy	168	Leu	Gly	Lys	Glu	Glu	Gly	Lys	Arg	Leu	Glu	Glu	His	Asp	Lys	Lys	Ile	Glu	Glu	187		
Db	481	TTAGG	AAA	GA	AAAA	GAAG	CTG	AGAT	TACT	TTA	AAA	AGT	ACG	ATG	ATA	AA	GAT	AGT	AGT	AGT	540	
Qy	188	Tyr	Lys	Lys	Glu	Ile	Thr	Met	Asp	Lys	Asn	Gln	Lys	Val	Leu	Pro	Ala	Val	Ala	Lys	207	
Db	541	TTCC	AAAA	---	---	---	GAT	GCA	AA	ACCA	AA	AGT	ATA	AA	GAT	GAT	GAT	GCC	ATT	GA	588	
Qy	208	Ser	Gly	Leu	Leu	Ala	His	Pro	Ser	Asn	-----	-----	-----	-----	Ser	Tyr	Val	Gly	220			
Db	589	GCT	TCA	GTT	GT	TAA	CTT	CCG	TGCT	GTAT	CAT	ACA	GA	AA	TTT	TAT	GCT	G	TG	TGAT	648	
Qy	221	Gln	Phe	Leu	Ser	Gln	Leu	Gly	Phe	Lys	Glu	Ala	Leu	Ser	Asp	Val	Thr	Lys	Gly	Leu	240	
Db	649	GAAT	CTT	TAAT	GAT	TT	TAG	GAT	TCAA	ACGT	-----	-----	-----	-----	AATA	AA	GACT	T	690			
Qy	241	Ser	Lys	Tyr	Leu	Lys	Gly	-----	Pro	Tyr	Leu	Gln	Met	Asn	Thr	-----	-----	-----	-----	-----	257	
Db	691	CAAAA	CA	AGT	TG	TAAT	TG	TAA	AGAT	TAT	T	CCA	CTT	CAT	CTA	AA	GA	AA	GA	AGT	750	
Qy	258	Gln	Val	Asn	Pro	Glu	Arg	Met	Phe	Ile	Met	Thr	Asn	Lys	Ala	Ser	Ser	Asn	Glu	Pro	Ser	277

RESULT 4

```

US-10-278-946-15
; Sequence 15, Application US/10278946
; Patent No. 6821754
; GENERAL INFORMATION:
; APPLICANT: Simpson et al.
; TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB461USb1
; CURRENT APPLICATION NUMBER: US/10/278,946
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: US 09/830,217
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: PCT/US99/06199
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: US 60/078,682
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 60/080,296
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: US 60/084,674
; PRIOR FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 999
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-278-946-15

```

Alignment Scores:		
Pred. No.:	7.57e-37	999
Score:	405.00	113
Percent Similarity:	51.74%	65
Best Local Similarity:	32.85%	132
Query Match:	23.96%	132
DB:	4	12
		Gaps: 3
		Indels: 14
		Mismatches: 132
		Conservative: 65
		Matches: 113
		Length: 999

US-10-724-972A-6352 (1-335) x US-10-278-946-15 (1-999)

QY		8	LeuLysIleLeuSerValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCysGly	27
Db		19	ATTAAATGCCTT--GGTTGTACGCTTGCTTCCTACTGCT--TTAGCAGGATCGACT	72
QY		28	AenAenSerSerAenSerLysGluSerSerLysAspGlyValGluIleLysHis	47
Db		73	CGGAATTCAAATAAACAAATCATCTGATACAAGAATAAGGAACAACCTTCAATTAAACAT	132
QY		48	GluGluGlyThrThrLysValProLysHisProLysArgValValValLuuGluTySer	67
Db		133	GCAATGGGTCAACTGAAATTAAAGGGAAACCACCGCTGTGTACGCTATATCAAGT	192
QY		68	PheValAspAlaLeuValAlaLeuAspValLysProValcglylLealaAspAspAenLys	87
Db		193	GCCACTGACGTCGCTGATCTTAGTGTTTAAACCTGTAGGTGCTGTGAANTCATGCACA	252

QY 88 LysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThr 107
 DB 253 CAAACCCGAAATTCGAATACATAAAAGATGATTAAAGATACTAAGATTGTAGGTCAA 312
 QY 108 ArgLysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIleIleAlaAsp 127
 DB 313 GAACCTGCACCTAACCTTAGAGGAATCTCTAAATTTAAACCGGACCTAAATGTGCGGTCA 372
 QY 128 AsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeu 147
 DB 373 AAAGTTAGAAATGAAAAAGTTTACGATCAATTAATCTAAATTCGCCAACCAAGT----- 426
 QY 148 LysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAla 167
 DB 427 ---TCTACTCATACAGTTTTCAAATTCAA--GATCAACTAAGTTAATGGGGAAAGCT 480
 QY 168 LeuGlyLysGluGluGluGlyLysLysArgLeuGluHisAspLysLysIleGluGlu 187
 DB 481 TTAGGGGAAGAAAAGAGCTGAAGATTCTTTAAAAAGTACGATGATAAAGTAGCTGCA 540
 QY 188 TyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaAlaLys 207
 DB 541 TTCCAAAA-----GATGCAAAAGCAAGTATAAAGATGCATGCCCATTTGAA 588
 QY 208 SerGlyLeuLeuAlaHisProSerAsn-----SerTyrValGly 220
 DB 589 GCTTCAGTTGTAACTCCGTCGTGATCATACAAGATTTTATGCTGGTGGATATGCTGGT 648
 QY 221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu 240
 DB 649 GAAATCTTAAATGATTTAGGATTCAAACGT-----AATAAAGACTTA 690
 QY 241 SerLysTyrLeuLysGly-----ProTyrLeuGlnMetAsnThr---GluThrLeuSer 257
 DB 691 CAAACCAAGTTGATTAATGGTAAAGATATTATCCAACTTACATCAATAAGAAAGCATTCGA 750
 QY 258 GlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSer 277
 DB 751 TTAATGAACGCTGATCATATTTTGTAGTAAATCAGATCCAAATGCGAAGATGCTGCA 810
 QY 278 Leu-----LysGluLeuGluLysAspProValTyrLysLysLeuAsnAla 292
 DB 811 TTAGTTAAAAAGACTGAAAGCGAATGGACTTCAAGTAAAGAGTGGAAAAAATTAGACGCA 870
 QY 293 ValLysAsnGlnArgVal---AspIleLeuAspArgAspLeuTyrAlaArgSerArgGly 311
 DB 871 GTTAAAAACAACCAAGTATCTCATGATTTAGATGAATCATCTTGGNACTTAGCTGGCGGA 930
 QY 312 LeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAspSerLys 331
 DB 931 TATAATCTTCATTAACACTTATTGACGATTTATATGAA-----AAGTTAAATATTGAA 984
 QY 332 LysAspAsnLys 335
 DB 985 AAACAATCAAA 996

RESULT 5
 US-08-956-171E-238
 ; Sequence 238, Application US/08956171E
 ; Patent No. 6593114
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; Gil H. Choi
 ; Patrick S. Dillon
 ; Craig A. Rosen
 ; Steven C. Barash
 ; Michael R. Fannon
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 5256
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS/EE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ;

```

STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1

TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 238:
SEQUENCE CHARACTERISTICS:
LENGTH: 3775 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 238:
US-08-956-171E-238

Alignment Scores:
Pred. No.: 5,58e-36 Length: 3775
Score: 405.00 Matches: 113
Percent Similarity: 51.74% Conservative: 65
Best Local Similarity: 32.85% Mismatches: 132
Query Match: 23.96% Indels: 34
DB: 4 Gaps: 12

US-10-724-972A-6352 (1-335) x US-08-956-171E-238 (1-3775)
QY 8 LeuValleLeuSerValleGlyLeuLeuPheValleuLeuAlaThrAlaAaCysGly 27
Db 61 ATTTAAATGCTT--GTGTAGCTTGCTTCTTCTACTTGT---TTAGCAGGATGTAGT 114
QY 28 AsnAsnSerSerAsnSerSerLysGluSerSerLysAspGlyValGluIleYehis 47
Db 115 GCGAATTCAAATAAACAATCATCTGTGTAACAAAGATAAGGAAACAACCTTCAATTAACACT 174
QY 48 GluGluGlyThrThrLysValProLysHisProLysArgValValValLeuGluTyrSer 67
Db 175 GCAATGGGTCAACTGAAATTTAAAGGGAACCAAGCGGTGTGTACCGTATATCAAGT 234
QY 68 PheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsnLys 87
Db 235 GCCACTGACGTGCGTGTATCTTTAGTGTGTTAAACCTGTAGTGCTGTAGAATCATCGACA 294
QY 88 LysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThr 107
Db 295 CAAAAACCGAAATTCGAATCATATAAAATGATTTAAAGATACCTAAGATTGTAGGTCAA 354
QY 108 ArgLysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIleLeuAsp 127
Db 355 GAACCTGCACCTTACTTAGAGGAATCTCTAAATTTAAACCGGACTTAATTTGCGGTCA 414
QY 128 AsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeu 147
Db 415 AAAGTTAGAAATGAAAAAGTTTACGATCAATTTCTAAATTCGCAACACAGTT----- 468
QY 148 LysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAla 167

```

```

RESULT 5
US-08-956-171E-238
; Sequence 238, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
;           Gil H. Choi
;           Patrick S. Dillon
;           Craig A. Rosen
;           Steven C. Barash
;           Michael R. Fannon
TITLE OF INVENTION: Staphylococci
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Human Genome Science Center
STREET: 9410 Key West Avenue
CITY: Rockville

```



```
Db 469 ---TCTACTGATACAGTCTTTTCAAAATTCAAA---GATACAACTAAGTTAATGGGAAAGCT 522
Qy 168 LeuGlyLysGluGluGluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGlu 187
Db 523 TTAGGGAAGAAAGAAAGAGCTGAAGATTTTACTTTAAAGAGTACGATGATAAGTAGCTGCA 582
Qy 188 TyrLysLysGluLysLeuThrMetAspLysAsnGlnLysValLeuProLysValAlaLys 207
Db 583 TTCCAAAAA-----GATGCAAAAGCAAGTATATAAGATGCGCCATTGAAA 630
Qy 208 SerGlyLeuLeuAlaHisProSerAsn-----SeryrValGly 220
Db 631 GCTTCAGTTGTTAACTTCGCTGCTGATCATACAGAAATTTATCTCGTGATATGCTGGT 690
Qy 221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu 240
Db 691 GAAATCTTAATGATTTAGGATTCAAACGT-----AATAAGACTTA 732
Qy 241 SerLysTyLeuLysGly-----ProTyLeuGlnMetAsnThr-----GluThrLeuSer 257
Db 733 CAAAAACAAAGTTGATAATAGTAAAGATATTATCCAACTTACATCTAAAGAAAGCATTC 792
Qy 258 GlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSer 277
Db 793 TTAATGACGCTGATCATATTTTGTAGTAAATCAGATCCAAATGCGAAAGATGCTGCA 852
Qy 278 Leu-----LysGluLeuGluLysAspProValTrpLysLysLeuAsnAla 292
Db 853 TTAGTTAAAAAGACTGAAAGCGAATGAGCTTCAAGTAAAGACTGGAATAATTAGACGCA 912
Qy 293 ValLysAsnGlnArgVal---AspIleLeuAspArgAspLeuTrpAlaArgSerArgGly 311
Db 913 GTTAAAAACAACCAAGTATCTGATGATTTAGATGAAATCAGCTTGAACCTTAGCTGGCGA 972
Qy 312 LeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAspSerLys 331
Db 973 TATAAATCTTCAATTAATTAATGACGATTTATATGAA-----AAGTTAAATATTGAA 1026
Qy 332 LysAspAsnLys 335
Db 1027 AAACAATCAAAA 1038

RESULT 6
US-08-781-986A-238
; Sequence 238, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781.986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 238:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3775 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-238
```

```
Alignment Scores:
Pred. No.: 5,58e-36 Length: 3775
Score: 405.00 Matches: 113
Percent Similarity: 51.74% Conservative: 65
Best Local Similarity: 32.85% Mismatches: 132
Query Match: 23.96% Indels: 34
DB: 4 Gaps: 12
```

US-10-724-972A-6352 (1-335) x US-08-781-986A-238 (1-3775)

```
Qy 8 LeuLysIleLeuSerValIleGlyLeuLeuPheValLeuAlaThrAlaLysGly 27
Db 61 ATTAATAATGCTT---GTTGTTACGCTTGTCTTCTTACTTGT---TTAGCAGGATGCTAGT 114
Qy 28 AsnAsnSerSerSerAsnSerSerLysGluSerSerLysAspGlyValGluLysHis 47
Db 115 GGGAAATTCAAATATAAACAATCATCTGATAACAAAGATAAGAAACAACATTTCAATTAACAT 174
Qy 48 GluGluGlyThrThrLysValProLysHisProLysArgValValValValLeuGluTySer 67
Db 175 GCAATGGGTCAACACTGAATTAAGCGAAACCAAGCGGTGTTGTACGCTATATCAAGT 234
Qy 68 PheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLys 87
Db 235 GCCACTGACGTCGCTGTATCTTTAGTGTATAAACCTGTAGGTGCTAGTAATCATGGACA 294
Qy 88 LysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyThrSerValGlyThr 107
Db 295 CAAAAACCGAAATTCGAATATCAATAAATAAGATTTAAAGATACTAAGATTGTAGTCAA 354
Qy 108 ArgLysGlnProAsnLeuGluLysSerLysLysLysProAspLeuIleIleAlaAsp 127
Db 355 GAACCTGCACCTAATTAGAGAAATCTCTAAATTAACACCGAGCTTAATTTGTCGCTCA 414
Qy 128 AsnAsnArgHisLysGlyIleTyLysAspLeuAsnLysIleAlaProThrIleGluLeu 147
Db 415 AAAGTTAGAATAAGAAAGTTTACGATCAATTAATTAATAATCGCACCAACAGTT----- 468
Qy 148 LysSerPheAspGlyAspTyAsnGluAsnIleAspAlaPheLysThrIleSerLysAla 167
Db 469 ---TCTACTGATACAGTTTTCAAATTCAAA---GATACAACTAAGTTAATGGGAAAGCT 522
Qy 168 LeuGlyLysGluGluGlyLysLysArgLeuGluHisAspLysLysIleGluGlu 187
Db 523 TTAGGGAAGAAAGAAAGAGCTGAAGATTTACTTAAAGATACGATGATAAGATGAGTCTCA 582
Qy 188 TyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLys 207
Db 583 TTCCAAAAA-----GATGCAAAAGCAAGTATATAAGATGCGATGCCATTGAAA 630
Qy 208 SerGlyLeuLeuAlaHisProSerAsn-----SeryrValGly 220
Db 631 GCTTCAGTTGTTAACTTCGCTGCTGATCATACAGAAATTTATGCTGGTATATGCTGGT 690
Qy 221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu 240
Db 691 GAAATCTTAATGATTTAGGATTCAAACGT-----AATAAGACTTA 732
Qy 241 SerLysTyLeuLysGly-----ProTyLeuGlnMetAsnThr-----GluThrLeuSer 257
Db 733 CAAAAACAAAGTTGATAATAGTAAAGATATTATCCAACTTACATCTAAAGAAAGCATTC 792
```

QY 258 GlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSer 277
DB 793 TTAATGAACGCTGATCATATTTTGTAGTAAATCAGATCCAAATCGGAAGATGCTGCA 852
QY 278 Leu-----LysGluLeuGluLysAspProValTrpLysLysLeuAsnAla 292
DB 853 TTAGTTAAAAGACTGAAAGCGAATCGACTTCAAGTAAAGAGTGGAAAAATTTAGACGCA 912
QY 293 ValLysAsnGlnArgVal---AspIleLeuAspArgAspLeuTrpAlaArgSerArgGly 311
DB 913 GTTAAAAACAACCAAGTATCTGATGATTTAGATGAAATCATTGGAACATTAGCTGCGGA 972
QY 312 LeuLysSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAspSerLys 331
DB 973 TATAAATCTTCATTAACACTTATTGAGATTATATGAA-----AAGTTAAATATGAA 1026
QY 332 LysAspAsnLys 335
DB 1027 AAACAATCAAAA 1038
RESULT 7
US-08-956-171E-604
; Sequence 604, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hymen
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 604:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2115 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 604:
US-08-956-171E-604
Alignment Scores: 3.35e-28 Length: 2115
Pred. No.: 334.00 Matches: 67
Score:

Percent Similarity: 63.84% Conservative: 46
Best Local Similarity: 37.83% Mismatches: 58
Query Match: 19.76% Indels: 6
DB: 4 Gaps: 3
US-10-724-972A-6352 (1-335) x US-08-956-171E-604 (1-2115)
QY 154 TyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGlu 173
DB 3 TATATGCAATATTTGAAGCATTTTAAACAGTCGCTTAAAGCAGTAGCGCAAGAGAAAGAA 62
QY 174 GlyLysLysArgLeuGluGluHisAspLysLysIleGluGluLysLysGluLysLys 193
DB 63 GCGGAGAACGCTCTGGAAGAGCATGATAAAATATTAGCCGAGATTAGAAAGAAATTTGAA 122
QY 194 MetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHis 213
DB 123 CAGAGTACGTTAAATCTGCATTTGCGGTATCTCAAGAGCAGGTATGTTTATTAAT 182
QY 214 ProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSer 233
DB 183 AATGAAGATACATTTATGGGACAAATTTCTTAAATTTAAATGGGTATTCAACCTGAAGTCAMA 242
QY 234 AspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThr 253
DB 243 AAARAAAAAACTACCGATGTTGGTGAACCGAAGGGTGTCTTATATATATTTAAATAAT 302
QY 254 GluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn---LysAlaSer 272
DB 303 GAAGAACTTCCCAATATCAATCCAAAGATTATGATTTTAGCCACTGACGAAAAACGGAC 362
QY 273 SerAsnGluProSerLeuLysGluLeuGlyLysAspPro---ValTrpLysLysLeuAsn 291
DB 363 AAAAAATAGAACGAAATTC-----ATTGATCTCGAGTTTGGAAATCATTAATAA 410
QY 292 AlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArgGly 311
DB 411 GCTGTGAAGATACAAAGATTTATGACGTGTACCCGAATAAGTGTGTTGAAATCAAGGGG 470
QY 312 LeuLysSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328
DB 471 ATTATCGCAAGTGAAGTATGGCAGAGATTTAGAAAAAATTTGACAGAAAA 521
RESULT 8
US-08-781-986A-604
; Sequence 604, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446

REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 604:
SEQUENCE CHARACTERISTICS:
LENGTH: 2115 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-604

Alignment Scores:
Pred. No.: 3,35e-28 Length: 2115
Score: 334.00 Matches: 67
Percent Similarity: 63.84% Conservative: 46
Best Local Similarity: 37.85% Mismatches: 58
Query Match: 19.76% Indels: 6
DB: 4 Gaps: 3

US-10-724-972A-6352 (1-335) x US-08-781-986A-604 (1-2115)

```
QY 154 TyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 173
Db 3 TATAATGCAATATTGAAGCATTTAAACAGTCGCTTAACAGTAGGCAAGAAGAA 62
QY 174 GlyLysLysArgLeuGluGluHisAspLysLysIleGluGlyLysGluIleThr 193
Db 63 GCGCAGAAGCGCTCGGAAGACGATATAAATATTAGCGAGATTAGAAAGAAATTGAA 122
QY 194 MetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHis 213
Db 123 CAGAGTAGCTTAATATCTGCAATTCGATTCGGTATCTCAAGACGAGTATGTTTATTAAT 182
QY 214 ProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSer 233
Db 183 AATGAAGATACATTTATGGACATTTCTTAATTAATGGGTATTCACCTGAAGTCAMA 242
QY 234 AspAspValThrLysGlyLysSerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThr 253
Db 243 AAARAAATAACTACGCATGTGGTGAACGCAAGGGTGGCTTATATATATATTAATAAT 302
QY 254 GluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn---LysAlaSer 272
Db 303 GAAGAACTTGGCAATATCAATCCAAAGATTATGATTTAGCCCACTGACGGAACCGGAC 362
QY 273 SerAsnGluProSerLeuLysGluLeuGlyLysAspPro---ValTyrLysLysLeuAsn 291
Db 363 AAAAATAGAACGAATTC-----ATTGATCCTGCAGTTTGGAAATCATTAATA 410
QY 292 AlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTyrAlaArgSerArgGly 311
Db 411 GCTGTGAAGATAACAAAGTTTATGACGTTGACCGAATAAGTGTGGAATCAAGGGGG 470
QY 312 LeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328
Db 471 ATTATCGCAAGTGAAGTATGCGAGAAGATTTAGAAAAAATTCAGAAAAA 521
```

RESULT 9

US-08-956-171E-2556
Sequence 2556, Application US/08956171E
Patent No. 6593114

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

Gil H. Choi

Patrick S. Dillon

Craig A. Rosen

Steven C. Barash

Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5255

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 2556:
SEQUENCE CHARACTERISTICS:
LENGTH: 242 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2556:
US-08-956-171E-2556

Alignment Scores:
Pred. No.: 1.68e-29 Length: 242
Score: 333.00 Matches: 66
Percent Similarity: 90.00% Conservative: 6
Best Local Similarity: 82.50% Mismatches: 8
Query Match: 19.70% Indels: 0
DB: 4 Gaps: 0
US-10-724-972A-6352 (1-335) x US-08-956-171E-2556 (1-242)

```
QY 70 AspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsnLysLysAsn 89
Db 2 GATGATTTAGCAGCATTAGACGTTAAACAGTTGGTATTGCTGATGATGCTAAGAAAAA 61
QY 90 ArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLys 109
Db 62 CGTATCATTAACACGATTAGAGAAAAAATGGGGATTATCTCTCTGTAGGTACAGCTAAA 121
QY 110 GlnProAsnLeuGluGluIleSerLysLysLeuLysProAspLeuIleIleAlaAspAsnAsn 129
Db 122 CAGCCAACTTAGAAGAAATTTAGTAATTAACCGGATTTAATATTCGCTGATAGCAGT 181
QY 130 ArgHisLysGlyIleTyrLysAspLeuAsnLysLysIleAlaProThrIleGluLeuLysSer 149
Db 182 AGACATANAGGTATTATAAAGAAATTAAACAAATTCACCAACATTATCATTAAGAGT 241
```

RESULT 10

US-08-781-986A-2556

Sequence 2556, Application US/08781986A

Patent No. 6737248

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5255

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

```
/
/
/ MEDIUM TYPE: CD/ROM ISO9660
/ COMPUTER: PC
/ OPERATING SYSTEM: <Unknown>
/ SOFTWARE: ASCII
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/107,532A
/ FILING DATE: 30-Jun-1998
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 60/085,598
/ FILING DATE: 14 May 1998
/ APPLICATION NUMBER: 60/051571
/ FILING DATE: July 2, 1997
/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Ariniello, Pamela Deneke
/ REGISTRATION NUMBER: 40,489
/ REFERENCE/DOCKET NUMBER: GTC-012
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (781)893-5007
/ TELEFAX: (781)893-8277
/
/ INFORMATION FOR SEQ ID NO: 862:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1059 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: circular
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Enterococcus faecium
/ FEATURE:
/
/ NAME/KEY: misc feature
/ LOCATION: (B) LOCATION 1...1059
/ SEQUENCE DESCRIPTION: SEQ ID NO: 862:
/
/ US-09-107-532A-862
/
/ Alignment Scores:
/ Pred. No.: 1.31e-25 Length: 1059
/ Score: 307.50 Matches: 107
/ Percent Similarity: 50.29% Conservative: 65
/ Best Local Similarity: 31.29% Mismatches: 123
/ Query Match: 19.70% Indels: 47
/ DB: 4 Gaps: 15
/
/ US-10-724-972A-6352 (1-335) x US-09-107-532A-862 (1-1059)
/
/ QY 5 ValarglyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeuIleAlaThrAla 24
/ DB 118 GTAGTAGGTATCTTGATTTCTAGCATCATTTGGTTT-----GCA 156
/
/ QY 25 AlaCysGly-----AsnAsnSerSerSerSerSerSerSerSerSerSerSer 39
/ DB 157 GCATGTGGAATAGCAACACACAGTCAGGCAGACACAAACAAAGACGACACAAACT 216
/
/ QY 40 LysAspGlyValGluIleLysHisGluGluGly-----ThrThrLysValProLysHisPro 58
/ DB 217 GAAACTACATTCAGCATTAACAGATAGTAAATCGAGACCAAAATCGAGGTGCCAAATAATCCA 276
/
/ QY 59 LysArgValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspVal--- 77
/ DB 277 GAAAGAGTCGTGATTCGACAAATGTTCTTAGATAGATGATGCTCTTGGTGTGGA 336
/
/ QY 78 ---LysProValGlyIleAlaAspAspAsnLysLysAsnArgIleIleLysProLeuArg 96
/ DB 337 GATAGGTAGTGGAGCAGCTACAGCAGC-----CTTCTGAGTATCTT 381
/
/ QY 97 AspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluIle 116
/ DB 382 TCATCTCTATAAAAAAGTAGAATCAGCAGGTGGAATAAAAGAACGACACTAGAAAAAT 441
/
/ QY 117 SerLysLeuLysProAspLeuIleAlaAspAsnAsnAsnHisLysGlyIleTyrLys 136
/ DB 442 AATCAGCTTCAGCCAGATCTAATCATCATTT---TCTGCCGACAAAGAGATTTTCAAAA 498
/
/
/ STATE: Maryland
/ COUNTRY: USA
/ ZIP: 20850
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
/ COMPUTER: HP Vectra 486/33
/ OPERATING SYSTEM: MSDOS version 6.2
/ SOFTWARE: ASCII Text
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/781,986A
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Benson, Bob
/ REGISTRATION NUMBER: 30,446
/ REFERENCE/DOCKET NUMBER: PB248PP
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (301) 309-8504
/ TELEFAX: (301) 309-8512
/
/ INFORMATION FOR SEQ ID NO: 2556:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 242 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/
/ US-08-781-986A-2556
/
/ Alignment Scores:
/ Pred. No.: 1.68e-29 Length: 242
/ Score: 333.00 Matches: 66
/ Percent Similarity: 90.00% Conservative: 6
/ Best Local Similarity: 82.50% Mismatches: 8
/ Query Match: 19.70% Indels: 0
/ DB: 4 Gaps: 0
/
/ US-10-724-972A-6352 (1-335) x US-08-781-986A-2556 (1-242)
/
/ QY 70 AspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsnLysLysAsn 89
/ DB 2 GAATCATTAAGCAGCATTAGACGTTAAACCGATTGGTATTGCTGATGGTAAAGAAAAA 61
/
/ QY 90 ArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLys 109
/ DB 62 CGTATCATTAACACGATTAGAGAAAAAATTGGGATTATCTCTGTAGGTACACGTAAA 121
/
/ QY 110 GlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsnAsn 129
/ DB 122 CAGCCAAACTTGAAGAAATTAGTAAATTAAACCGGATTTTATTCGCTGATGACAGT 181
/
/ QY 130 ArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSer 149
/ DB 182 AGACATANAGGTATTAAATAAGATTAAACAAATTAACAAATTAACCAACATTATCAATTAAGAGT 241
/
/ RESULT 11
/ US-09-107-532A-862
/ Sequence 862, Application US/09107532A
/ Patent No. 6583275
/
/ GENERAL INFORMATION:
/ APPLICANT: Lynn A Doucette-Stamm and David Bush
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
/ ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
/
/ NUMBER OF SEQUENCES: 7310
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: GENOME THERAPEUTICS CORPORATION
/ STREET: 100 Beaver Street
/ CITY: Waltham
/ STATE: Massachusetts
/ COUNTRY: USA
/ ZIP: 02354
/
/ COMPUTER READABLE FORM:
```


Db	961	GATCCAAACCTATGGTATTCTCTTCAGGA-----TCTTCAACGACCAACTATCAACAA	1014
Qy	322	LeuValGluLeuSerLys 327	
Db	1015	ATTGATGAATTAATGAA 1032	
RESULT 13			
US-09-710-279-3348/c			
; Sequence 3348, Application US/09710279			
; Patent No. 6703492			
; GENERAL INFORMATION:			
; APPLICANT: KIMMERLY, WILLIAM JOHN			
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS			
; FILE REFERENCE: PUS480US			
; CURRENT FILING DATE: 2000-11-09			
; PRIOR APPLICATION NUMBER: 60/164,258			
; PRIOR FILING DATE: 1999-11-09			
; NUMBER OF SEQ ID NOS: 4472			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 3348			
; TYPE: DNA			
; ORGANISM: Artificial Sequence			
; FEATURES:			
; OTHER INFORMATION: Description of Artificial Sequence: synthetic			
; OTHER INFORMATION: nucleic acid sequence			
US-09-710-279-3348			
Alignment Scores:			
Pred. No.: 1,248-23 Length: 3153			
Score: 296.50 Matches: 104			
Percent Similarity: 47.54% Conservatives: 70			
Best Local Similarity: 28.42% Mismatches: 131			
Query Match: 17.54% Indels: 61			
DB: 4 Gaps: 16			
US-10-724-972A-6352 (1-335) x US-09-710-279-3348 (1-3153)			
Qy	1	GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeu 20	
Db	1299	GGAGTTGCACATATGAAGAAAACAGTCTTATTTTATTTATTTATTTCTCTAGTTTATTTA 1240	
Qy	21	IleAlaThrAlaAlaCysGlyAsnAsnSerSerSerAsnSerSerLysGluSer---Ser 39	
Db	1239	ACGGCTTGTAAGTATAGTTTCGAATAATATTCACACTTCGAAAAAGAAAAATAGTATCT 1180	
Qy	40	LysAspGlyValGluIleLysHisGlu----- 48	
Db	1179	NAAGAACTGTAAACCATCAAAATAGTTTGAAGCAAGTGTGAAGAAAATATGCGACT 1120	
Qy	49	-----GluGlyThrThrLysValProLysHisProLysArgValVal 63	
Db	1119	GATAGAAAAAATCTCTAATACTGTCGAAGTACCAAGAACTCTAAAAATCGCGTTGTA 1060	
Qy	64	LeuGluTyrSerPheValAspAlaLeuAlaLeuAspValLysProValGlyIleAla 83	
Db	1059	TTAGATTATGAGCGCTTGATGTGTG-----AAAGAATTAGGTGGCT 1015	
Qy	84	AspAspAsnLys-----LysAsnArgIleIleLysProLeuArgAsp 97	
Db	1014	GATAAAGTAAAGGTTTACCTAAAGTGAAATACCAATCTTTACTTAAATTTTAGAT 955	
Qy	98	-----LysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsnLeuGlu 115	
Db	954	GAATTTAAAGATGATAGTATATTAATCTGGAATTTTAAAGAAAGTGAACCTTTGATAA 895	
Qy	116	IleSerLysLeuLysProAspLeuIle---IleAlaAspAsnAsnArgHisLysGlyIle 134	
Db	894	GTTGCATCAGCTAAACCAAGATGTATTTTATTTTACAGGAAGAACAGCTAATCAGAAAAAT 835	
Qy	135	TyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGlyAspTyr 154	

Db	834	TTAGATGAATTTAAAAAGCTGCACCAAAAGCTAAAGTT---GTATATGTAGTACAAGT 778	
Qy	155	AsnGluAsnIle-----AspAlaPheLys-----ThrIleSerLysAlaLeuGly 169	
Db	777	GATGACAACTTAATTAAGATATGAAAAAATACAGAAAAATTTAGGAAAAATCTACGAT 718	
Qy	170	LysGluGluGlyLysArgLeuGluGluHisAspLysLysIleGluGluTyrLys 189	
Db	717	AAAGAAGATAAGCTAAAAAATAAATAAGATTTAGATAGAAAAATATCTGATATGAAA 658	
Qy	190	LysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGly 209	
Db	657	GATAAA---ACTAAGACTTTAATAAGAAAGTAAATGATTATTATTGTTAAGAGGTTGAA 601	
Qy	210	LeuLeuAlaHis---ProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPhe 228	
Db	600	CTATCAACGTTTGGACCAAGGAGGATTTGGTGGTTAGTTGTTGATACATTAGGATTT 541	
Qy	229	LysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyr 248	
Db	540	AAACCTGCA-----CACAAAAAGGTTAGCAAAAAAGCCCGCAT 505	
Qy	249	LeuGln---MetAsnThrGluThrLeuSerGlnValAsnPro----- 261	
Db	504	GGTCAAAATATAAATAATGAATATTAACAAGCAGAAATCCAGATGTTATTTTAGCTATG 445	
Qy	262	GluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu 281	
Db	444	GATCGTGTTCAGTTGTTAGTGGTTAAGCAACAACAAAT-----CAAGTT 400	
Qy	282	GluLysAspProValTrpLysLeuAsnAlaValLysAsnGlnArgValAspIleLeu 301	
Db	399	TTAAAAACAAGTTATATAAAAAATGTAAGCAGTAAAAAGTAAATCATATTTTACGAATTA 340	
Qy	302	AspArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGlu 321	
Db	339	GATCCAAACTATGTTCTTCTTCAGGA-----TCTTCAACGACAACTATCAACAA 286	
Qy	322	LeuValGluLeuSerLys 327	
Db	285	ATTGATGAATTAATGAA 268	

RESULT 14

US-09-710-279-4042

; Sequence 4042, Application US/09710279

; Patent No. 6703492

; GENERAL INFORMATION:

; APPLICANT: KIMMERLY, WILLIAM JOHN

; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS

; FILE REFERENCE: PUS480US

; CURRENT APPLICATION NUMBER: US/09/710,279

; CURRENT FILING DATE: 2000-11-09

; PRIOR APPLICATION NUMBER: 60/164,258

; PRIOR FILING DATE: 1999-11-09

; NUMBER OF SEQ ID NOS: 4472

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4042

; LENGTH: 3267

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURES:

; OTHER INFORMATION: Description of Artificial Sequence: synthetic

; OTHER INFORMATION: nucleic acid sequence

US-09-710-279-4042

Alignment Scores:

Pred. No.: 1,316-23 Length: 3267

Score: 296.50 Matches: 104

Percent Similarity: 47.54% Conservatives: 70

Best Local Similarity: 28.42% Mismatches: 131

Query Match: 17.54% Indels: 61

DB: 4 Gaps: 16

[illegible]

1020 TTATGATGAATTTAAAAAGAGCTGCACCAAAAGCTAAAGTT---GTTATATGTAGGTACAAAGT 1076
 155 AsnGluAenIle-----AspAlaPhelys-----ThrIleSerLysAlaLeuGly 169
 1077 GATGCAACTTAATTAAGAGATATGAAAAAATAATACAGAAAAATTTAGGAAAAATCTACGAT 1136
 170 LysGluGluGluGluGlyLysArgLeuGluGluHisAspLysLysLysLysLysLysLys 189
 1137 AAAGAAGATGAAAGCTAAAGAAATTAATAAGAGATTTAGATAGAAAAATATCTGATATGAAA 1196
 190 LysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGly 209
 1197 GATAAA---ACTAAAGACTTTAATAAGAAAGTAATGATTTATTTGGTTAAAGAGGTGAA 1253
 210 LeuLeuAlaHis---ProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPhe 228
 1254 CTATCAACCTTTGGACCGAGGAAGATTTGGTGGTGTAGTTAGTTATGATACATTAGGATTT 1313
 229 LysGluAlaLeuSerAspValThrLysGlyLeuSerLysLysLysLysLysLysLys 248
 1314 AAACCTGCA-----GACAAAGAGTTAGCAAAAGCCCGAT 1349
 249 LeuGln---MetAsnThrGluThrLeuSerGlnValAsnPro----- 261
 1350 GGTCAAAATAATAATAATGAATATTAACAGCAGATCCAGATGTTATTTAGCTATG 1409
 262 GluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLysLysLysLys 281
 1410 GATCGTGGTTCAGTTCTAGTGGTAAAGCAACAAAGAT-----CAAGTT 1454
 282 GluLysAspProValTrpLysLysLysLeuAlaValLysAsnGlnArgValAspIleLeu 301
 1455 TTAATAAACCAAGATTTATAAAATGTAAGAGCAGTAAAGATTAATCATATTTACGAATTA 1514
 302 AspArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGlu 321
 1515 GATCCAAACTATGGTATTCCTCTTCAGCA-----TCTTCAACGACAACTATCAACAA 1568
 322 LeuValGluLeuSerLys 327
 1569 ATTGATGAATTAATGAA 1586
 RESULT 16
 US-09-543-681A-2690
 ; Sequence 2690, Application US/09543681A
 ; Patent No. 6605709
 ; GENERAL INFORMATION:
 ; APPLICANT: GARY BRETON
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 ; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/543.681A
 ; PRIOR FILING DATE: 2000-04-05
 ; PRIOR APPLICATION NUMBER: US 60/128,706
 ; PRIOR FILING DATE: 1999-04-09
 ; NUMBER OF SEQ ID NOS: 8344
 ; SEQ ID NO 2690
 ; LENGTH: 1056
 ; TYPE: DNA
 ; ORGANISM: Proteus mirabilis
 US-09-543-681A-2690
 Alignment Scores:
 Pred. No.: 2,73e-24 Length: 1056
 Score: 296.00 Matches: 90
 Percent Similarity: 50.65% Conservative: 66
 Best Local Similarity: 29.22% Mismatches: 118
 Query Match: 17.51% Indels: 34
 DB: 11 Gaps: 11
 US-10-724-972A-6352 (1-335) x US-09-543-681A-2690 (1-1056)
 100 AAATCAATTAAGCCCACTCTCTTATTTGCTCTCTCTGTTATTCGCAGGGTTGTGATAC 165
 29 AsnSerSerSerAsnSerSerLysGluSerLysAspGlyValGluIleLysHisGlu 48
 166 GCACAAGATACCTCAACACCGAATCCACAGAGAAACAACTCTCTCACTATCGAACATGCT 225
 49 GluGlyThrThrLysValProLysHisProLysArgValValValLeuLeuGluTyrSerPhe 68
 226 CAAGCACCACCTGATGCTCTCAACCAAAAGTGGTTGTGATGAACATCGAACA 285
 69 ValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLysLys 88
 286 CTTGATATTGTTGATGCTCTTTGGCGTACCTGTTGTTGGCCCTACCAACAAACATCGTCAC 345
 89 AsnArgIleLysProLeuArgAspLysLysLysLysLysLysLysLysLysLysLys 103
 346 -----CTACCTAAATTTCTAGAGAATATACCAACAGAAATGAGTAC 387
 104 ---SerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLysLysLysLys 122
 388 ATCAACGAAGGTGGCTGTTTCAACCAAACTACGAAAAAATCAGTACTACTGCTCCCTGAT 447
 123 LeuIleAlaAspAsnAsnArgHisLysGlyIleLysLysLysLysLysLysLysLysLys 142
 448 CTGATTTTAAACG---GGTAGTCGTCTCGCATGATATGCCAAATTAAGTGAAATCGCA 504
 143 ProThrIleLysLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPhe--- 161
 505 CCATCCATT-----TCTATGATATCGACAGCACACGCTTTATTGACAGCTCTACT 555
 162 -----LysThrIleSerLysAlaLeuGlyLysGluGluGluGluGlyLysLysLysLys 178
 556 GAGCGTACACGACTTTAGTCAAAATTTTGGTAAAGAAAGAAAGCAAGCTAAAAAATTATTA 615
 179 GluGluHisAspLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 198
 616 GCTGACTTCAATAGCAAAATTTGATACGCGTAAAGCAAA---ACCCAGATCGCGGT--- 669
 199 LysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyr 218
 670 AAAGCAATGCTGATTTAGTCAGTGGCGGTAATAATTTCCGCTTATGGCCAGCTCTCGC 729
 219 ValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAspValThrLys 238
 730 TTTGGTTTATCTACGATGTTTGGCTTTGAACTGATATATCTTCAT---AGCCCA 786
 239 GlyLeuSerLysTyrLysLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGln 258
 787 GGCTCTCACGGTAATATT-----GTTAACTCAGAACTACTGCTAAAA 828
 259 ValAsnProGluArgMetPheIleMetThrAsnLysAlaSer-----SerAsn 274
 829 CTTAATCCAGCTGGATGTTTGATTCATTGACCGTATGCGCAATTTGCTGTAAGATTCA 888
 275 GluProSerLysLysGluLeuGlyAspProValTrpLysLysLysLysLysLysLysLys 294
 889 CAACCTGCA---AAACAAAGCTTGATAATGCCCTAGTGTAGAAAGGTTTATGATGAAAT 945
 295 AsnGlnArgValAspIleLeuAsp 302
 946 AAAGACCAAAATTTATCTATCTGAT 969
 RESULT 17
 US-09-710-279-461
 ; Sequence 461, Application US/09710279
 ; Patent No. 6703492
 ; GENERAL INFORMATION:
 ; APPLICANT: KIMMERLY, WILLIAM JOHN
 ; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
 ; FILE REFERENCE: PUS480US
 ; CURRENT APPLICATION NUMBER: US/09/710,279
 ; CURRENT FILING DATE: 2000-11-09

```

; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 461
; LENGTH: 1044
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-461

Alignment Scores:
Pred. No.: 7,74e-24 Length: 1044
Score: 292.00 Matches: 101
Percent Similarity: 47.74% Conservative: 68
Best Local Similarity: 28.53% Mismatches: 121
Query Match: 17.28% Indels: 64
DB: 4 Gaps: 16

US-10-724-972A-6352 (1-335) x US-09-710-279-461 (1-1044)
QY 16 LeuLeuPheValLeuLeuAlaThrAlaAlaCysGlyAsnAsnSerSerSerSerSerSer 35
DB 25 TTATGTCTCTAGTTTGTAGTTTAAACGGCTTGTAGTAATAGTTTCAATAATAATCAACT 84
QY 36 LysGlu-----SerSerLysAspGlyValGluLeuLysHisGlu-----48
DB 85 TCGAAAAGAAAATAAGTAGTCTTAAAGAACTGTACCATCAAAATAATAGTTTGAAGCA 144
QY 49 -----GluglyThrThrLysValPro 55
DB 145 AGTGGTAAAGAAAATAATAGCAGTGATAGAAAATAATCTCTAATAGTCTGCAAGTACCA 204
QY 56 LysHisProLysArgValValValLeuGluTySerPheValAspAlaLeuValAlaLeu 75
DB 205 AGAATCCTTAAATAATCCGCTGTGATTAGATTATGAGCGCTGATGTTG-----255
QY 76 AspValLysProValGlyLeuAlaAspAspAsnLys-----LysAsn 89
DB 256 -----AAGAATTAGTGTGCTGATAAGTAAAGTAAAGGTTTACCTAAAGGTGAAAATAAC 309
QY 90 ArgLysLeuProLysArgAsp-----LysLysGlyLysTyThrSerValGlyThr 107
DB 310 CAATCTTACCTAAATTTTATAGTGAATTTAAAGATGATAGTATATTAATCTGGAAT 369
QY 108 ArgLysGlnProAsnLeuGluGluLeuSerLysLysLeuProAspLeuLeu-----126
DB 370 TTAAGAAGAGTGAAGTCTTGTATAAGTGTGCATCAGCTAAACAGATGCTGATTTTATTC 429
QY 127 AspAsnAsnArgHisLysGlyLysLysGlyLysAspLeuAsnLysLeuAlaProThrLeu 146
DB 430 CGAAGNACAGCTAATCAGAAAATTTAGATGAATTTAAAGAGCTGCACCAAAAGCTAAA 489
QY 147 LeuLysSerPheAspGlyAspTyThrAsnGluAsnLeu-----AspAlaPheLys 162
DB 490 GTT---GTATATGTAGTACAGTGTACACACTTAATTAAGATATGAAAAAATAACA 546
QY 163 ---ThrLysLeuAlaLeuGlyLysGluGluGluGlyLysLysArgLeuGluHis 181
DB 547 GAAATTTAGGGAAATCTTACGATAAAGAAAGATAAAGCTTAAATAAATAAGATTTA 606
QY 182 AspyLysLysLeuGluGluTyLysLysGluLeuThrMetAspLysAsnGlnLysValLeu 201
DB 607 GATGAAAATATCTGATGATGAAGATATAA---ACTAAAGACTTTTAATAAGAAAGTAAATG 663
QY 202 ProAlaValAlaAlaLysSerGlyLeuLeuAlaHis---ProSerAsnSerTyValGly 220
DB 664 TATTATTGGTTTAAAGAGGTGAATCATCAAGCTTTGGACCAAGGAGGAGGATTTGGTGGT 723
QY 221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu 240

```

```

RESULT 18
US-09-710-279-1267
; Sequence 1267, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1267
; LENGTH: 1044
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-1267

Alignment Scores:
Pred. No.: 7,74e-24 Length: 1044
Score: 292.00 Matches: 101
Percent Similarity: 47.74% Conservative: 68
Best Local Similarity: 28.53% Mismatches: 121
Query Match: 17.28% Indels: 64
DB: 4 Gaps: 16

US-10-724-972A-6352 (1-335) x US-09-710-279-1267 (1-1044)
QY 16 LeuLeuPheValLeuLeuAlaThrAlaAlaCysGlyAsnAsnSerSerSerSerSerSer 35
DB 25 TTATGTCTCTAGTTTGTAGTTTAAACGGCTTGTAGTAATAGTTTCAATAATAATCAACT 84
QY 36 LysGlu-----SerSerLysAspGlyValGluLeuLysHisGlu-----48
DB 85 TCGAAAAGAAAATAAGTAGTCTTAAAGAACTGTACCATCAAAATAATAGTTTGAAGCA 144
QY 49 -----GluglyThrThrLysValPro 55
DB 145 AGTGGTAAAGAAAATAATAGCAGTGATAGAAAATAATCTCTAATAGTCTGCAAGTACCA 204
QY 56 LysHisProLysArgValValValLeuGluTySerPheValAspAlaLeuValAlaLeu 75
DB 205 AGAATCCTTAAATAATCCGCTGTGATTAGATTATGAGCGCTGATGTTG-----255
QY 76 AspValLysProValGlyLeuAlaAspAspAsnLys-----LysAsn 89

```

ORGANISM: Myxococcus xanthus
US-09-902-540-2208

Alignment Scores:
Pred. No.: 5,55e-24 Length: 702
Score: 291.00 Matches: 70
Percent Similarity: 51.95% Conservative: 50
Best Local Similarity: 30.30% Mismatches: 93
Query Match: 17.22% Indels: 18
DB: 4 Gaps: 5

US-10-724-972A-6352 (1-335) x US-09-902-540-2208 (1-702)

QY 104 SerValGlyThrArgLysGlnProAsnLeuGluGluLeuSerLysLysLeuProAspLeu 123
Db 46 AGTGTGGTACAGAAAGCGAGCAAGCTTAGAAGCAATCGCTGCTTTAAAGCCTGACTTG 105

QY 124 IleIleAlaAspAsnLeuArgHisLysGlyIleTyrLysAspLeuLeuAlaPro 143
Db 106 ATTATCGAAATATAATGCGCAAGAAAGATATACGATCAGCTTAGTCAAAATGCTCCG 165

QY 144 ThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThr 163
Db 166 ACTGTATTCTCTGAGGAGCTGCGCGAGATGGAAATCAAACTTTAAGCTA 216

QY 164 IleSerLysAlaLeuGlyLysGluGluGlyLysArgLeuGluGluHisAspLys 183
Db 217 TATGCAAAAGCTGTAAACAAAGAAAGTAAAGAAAGTATAGCTGACTATGACAAT 276

QY 184 LysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAla 203
Db 277 CGCTAGCAGATTAAAGAAAGCGCTTGGAGATCAGTTAAACAAAGAAATCTCAGTTGTA 336

QY 204 ValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeu 223
Db 337 CGCTTTTACAGCTGAGAGCTTCGTATCTATCATATAAGATTCATCTCAGGTGTTATTTA 396

QY 224 SerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyr 243
Db 397 GATCAGCTTGGATTTGCTCGCCCTGAGTCTCAAGAC----- 432

QY 244 LeuLysGlyProTyrLeuGlnMetAsnThr-----GluThrLeuSerGlnValAsnPro 261
Db 433 ---AAGGATGATTTTGTCTGAATGAATGCAACAAAGAACCGCATTCGCGCAATGGACGCG 489

QY 262 GluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLeuGluLeu 281
Db 490 GATCAGCTTGGATTTTGTCTCTTATGAACGCGGAGCGGTGAAGCAACAAAGCTTCAAAAA 549

QY 282 Glu-----LysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAsp 299
Db 550 GAAATGATAACGATCCTCTCTTTAAAGCTAAAAGTGGCTCAAGCAACAAAGCTTCAAT 609

QY 300 IleLeuAspArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAla 319
Db 610 AAAGTAGATGATGCTACTTGGGATACAGCAGCGGTGTTAGCAGCAATAATCTGACTT 669

QY 320 LysGluLeu-----ValGluLeuSerLysLys 328
Db 670 GATGATATTGAAAAAATCTTCTTAGATAAAAAA 702

Db 256 -----AAGAATAGGTGGCTGATATAAGTAAAGGTTTACTTAAAGGTGAATAAC 309

QY 90 ArgIleLeuLysProLeuArgAsp-----LysIleGlyLysTyrThrSerValGlyThr 107
Db 310 CAATCTTTACCTAAATTTTAGATGAATTTAAAGATGATAGTATATTAATCTGGAAT 369

QY 108 ArgLysGlnProAsnLeuGluGluLeuSerLysLysLeuProAspLeuLeu-----IleAla 126
Db 370 TTAAGAAGAGTAACTTTGATAAAGTTGCACTACGCTAAACAGATGATTTTATTTCA 429

QY 127 AspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGlu 146
Db 430 GGAAGACAGCTAATCAGAAAAATTTAGATGAATTTAAAGAGCTGACCAAAAAAGCTAAA 489

QY 147 LeuLysSerPheAspGlyAspTyrAsnGluAsnIle-----AspAlaPheLys----- 162
Db 490 GTT---GTATATGTAGGTACAAAGTATGACAACTTAATTAAGATATATCAAAAAATACA 546

QY 163 ---ThrIleSerLysAlaLeuGlyLysGluGluGluGlyLysArgLeuGluGluHis 181
Db 547 GAAATTTAGGGAATACTACGATAAAGAGATTAAGCTAAAGAAATTAATAAGATTTA 606

QY 182 AspLysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeu 201
Db 607 GATGAAAAATATCTGATATGAAGATAA---ACTAAAGACTTTTAATAAGAAAGTAAAG 663

QY 202 ProAlaValAlaLysSerGlyLeuLeuAlaHis---ProSerAsnSerTyrValGly 220
Db 664 TATTATTGGTTAAGCAAGGTGAACACTATCAACGTTTGGACGAGGAGGAAGATTTGGTGT 723

QY 221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu 240
Db 724 TTAGTGTGTTGATACATTAGGATTTAAACCTGCA-----GAC 759

QY 241 SerLysTyrLeuLysGlyProTyrLeuGln---MetAsnThrGluThrLeuSerGlnVal 259
Db 760 AAAAGGTTAGCAAAAGCCGATCGTCAATAATAATAATGAATATTAACAAGCAG 819

QY 260 AsnPro-----GluArgMetPheIleMetThrAsnLysAlaSerSer 273
Db 820 AATCCAGATGTTATTTTAGCTATGATCGTGTTCAGTTGTAGGTGGTAAAGCAACA 879

QY 274 AsnGluProSerLeuLysGluLeuGluLysAspProValTrpLysLysLeuAsnAlaVal 293
Db 880 AAT-----CAAGTTTAAAAAACAAGTTTATAAAATGTAATAAGCAGTA 924

QY 294 LysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArgGlyLeuIle 313
Db 925 AAAGTAATCATATTAGCAATTAGATCCAAACTATGTTATTTCTTTCAGGA----- 978

QY 314 SerSerGluGluMetAlaLysGluLeuValGluLeuSerLys 327
Db 979 TCTTCAACGCAACTATCAACAAATTTGATGAATTAATGAA 1020

RESULT 20
US-09-902-540-1583
; Sequence 1583 Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540

RESULT 19
US-09-902-540-2208
; Sequence 2208 Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 2208
; LENGTH: 702
; TYPE: DNA


```

385 ATT---TCTGGTCGTCACCAAGATTATCAAGAAACAATTAAGGCAATTTGGCCCAACCAATT 441
146 GluLeu-----LysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPhe 161
442 TACTTACGCTAGATGCGCAAAAATCTTTGGGCAATCAACGAAACAATAATATCGAAACGCTTA 501
162 LysThrIleSerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHis 181
502 GGCACCTATT-----TTTGATAAAGAGAGGTAGCTAAAGAAAAAATAAATCGGCTTA 552
182 AspLysLysIleGluGluTyrLysLysGluGluLeuThrMetAspLysAsnGluLysValLeu 201
553 GAAAAAGAAATTTGCTGACGTGAAAAAACAAGCAGAGCTAGCGCGAAT---AATGCGCTT 609
202 ProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGln 221
610 GTTGTGTAGTTAAGCAAGGCAACTTTCCGCTTACGGAAGAGCTCTCGTTTCGGTTTA 669
222 PheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAsp-----ValThr 237
670 ATTCATGATACATTTGGCTTCAAGCAGCA-----GACGATAAGATTGAAGCTTCCACT 723
238 LysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSer 257
724 CATGGCAAACT-----GTTTCTTACGAATATGTTT 756
258 GlnValAsnProGluArgMetPheIleMet---ThrAsnLysAlaSerSerAsnGluPro 276
757 GAAAAAATCTCTGGGATTTCTTTTGGTAGATCGCACCAAGCAATTTGGTGGCAGCAT 816
277 SerLeuLysGluLeuGluLysAspProValTyrLysLysLeuAsnAlaValLysAsnGln 296
817 TCAAAAGATAAGCTCGCTGCAACGAATGATTCAAAAACCGATGCTGTGTAATAATGAT 876
297 ArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGlu 316
877 AAGTCATTATGCTTCAACCAAGATGTTGTTATCTAAGCGGTGGTGGATTAGAATCAATG 936
317 GluMetAlaLysGluLeuValGluLeuSerLysLys 328
937 CATTTG-----ATGATAGAAGATGTTAAAAAA 963

RESULT 22
US-09-134-000C-937
; Sequence 937, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 937
; LENGTH: 984
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-937

Alignment Scores:
Pred. No.: 7,66e-23 Length: 984
Score: 283.00 Matches: 101
Percent Similarity: 48.19% Conservative: 59
Best Local Similarity: 30.42% Mismatches: 130
Query Match: 16.75% Indels: 42
DB: 4 Gaps: 13

US-10-724-972A-6352 (1-335) x US-09-134-000C-937 (1-984)

```

RESULT 23
US-09-543-681A-1393
; Sequence 1393, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:

NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 889 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-071-035-39

Alignment Scores:
Pred. No.: 3,47e-21 Length: 889
Score: 268.00 Matches: 96
Percent Similarity: 47.32% Conservative: 54
Best Local Similarity: 30.28% Mismatches: 129
Query Match: 15.86% Indels: 38
DB: 3 Gaps: 12

US-10-724-972A-6352 (1-335) x US-09-071-035-39 (1-889)

QY 26 CysGlyAsnAsn-----SerSerSerAsnSerSerLysGluSerLysAspGly 42
Db 2 TGTCAAAACAAATAAAAAACAGCAGATCTGCACAAACAGAAACACACAGCTAAACCGAA 61
QY 43 ValGluLeuLysHisGluGluGlyThrThrLysValProLysHisProLysArgValVal 62
Db 62 GTCAAGTCAAGACCAACCAATGGTCAATTAACCGTCTCCCAAAATCTTAAGAAAGTCGT 121
QY 63 ValLeuGluThrSerPheValAlaLeuValAlaLeuAspVal-----LysProVal 80
Db 122 GTTTTGTGATAATGTTCTTGGATACAAATGATGCACTAGGTGTCGGTGACCGGTGTA 181
QY 81 GlyLeuAlaAspAspAsnLysLysAsnArgLysLysProLeuArgAspLysLysGly 100
Db 182 GGTGCGCAACTAAAAAT-----ATCCCTGCGGTATTTGAAAAAATATACCA 226
QY 101 LysThrThrSerValGlyThrArgLysGlnProAsnLeuGluGluLysSerLysLys 120
Db 227 AAAGTTGAATCAGCAGCGCGCATTAAGAACACCATTTAGAAAAAATCAATCACTAA 286
QY 121 ProAspLeuLeuAlaAspAsnArgHisLysGlyLysLysLysLysLysLys 140
Db 287 CCAGACTTAATTAATTATT---TCTGTCGTCAACAGAGATTATCAAGAAACAATTAAG 343
QY 141 IleAlaProThrIleGluLeu-----LysSerPheAspGlyAspTyrAsnGlu 156
Db 344 ATTGCGCAACCATTTACTTAGCTGTAGATGCAAAAAATCCTTGGGCATCAACGAAAC 403
QY 157 AsnIleAspAlaPheLysThrLysLysAlaLeuGlyLysGluGluGluGlyLysLys 176
Db 404 AATATCGAAACGTTAGGCATATT-----TTTGATAAAGAGAGGTAGCTAAAGAA 454
QY 177 ArgLeuGluGluHisAspLysLysLysLysLysLysLysLysLysLysLysLys 196
Db 455 AAAATACTGGCTTAGAAAAAGAAATTTGCTGACGTGAAAAAACAACAGCAGAACCT 514
QY 197 AsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsn 216
Db 515 AAT---AATGCGCTTGTGTGTAGTTAACGAAGGACAACTTTCCTTCGTAAGAGGC 571
QY 217 SerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAsp 235
Db 572 TCTCGTTTCGGTTTAAATTCATGATCATTTGGCTTCAAGACGCA-----GACGATA 625
QY 236 -----ValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsn 252
Db 626 ATTGAAGCTTCCACTCATCGGCAAACT-----GTTTCT 658
QY 253 ThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMet---ThrAsnLysAla 271

124 IleAlaAlaAspAsnArgHisLysGlyLysGlyLysLysLysLysLysLysLysLys 143
Db 355 ATCTTGTCTTCAACAAAGTTTCGGAAGAAACCGCTGCAAAAAATCAGCAGCAGCAGC 414
QY 144 ThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThr 163
Db 415 ACATCCCACTTCTCTATCTCTTCAACTGGAGGAAACATG-----ATGCTT 465
QY 164 IleSerLysAlaLeuGlyLysGluGluGlyLysLysLysLysLysLysLysLysLys 183
Db 466 CTTGCCAGCTGAGTGGAAAGAGAGAAAGCAAGAAATTTATTCAGACTATGAACAG 525
QY 184 LysIleGluGluThrLysLysGluLeuThr---MetAspLysAsnGlnLysValLeuPro 202
Db 526 GATCTAAAGAAATATAAAACAAATAACACGATAAACGGAAGATTCAAAAGCCTTGT 585
QY 203 AlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe 222
Db 586 ATCAGATCAGACAGCAACATTTACCTTACCTGACAGGTTATTTCAACTCCACA 645
QY 223 Leu---SerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSer 241
Db 646 CTATACGGTGTATTAGGCTTTAAG-----CGCGCGAACGAAGTAAAGGCTGCAAAAGCG 699
QY 242 LysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnPro 261
Db 700 CAAGAGCTGAGT-----TCATTAGAAAAATTAAGTGAATGAAACCCG 741
QY 262 GluArgMetPheIle---MetThrAsnLysAlaSerSerAsnGluPro---SerLeuLys 279
Db 742 GACCATATTTTCGTCCTCAATTTCTGATGATGAAATGCAACACCTGATGCTTAA 801
QY 280 GluLeuGluLysAspProValTrpLysLysLysLysLysLysLysLysLysLysLys 299
Db 802 GATTTAGAGAAAAATCCAATCTGGAAGAGCCTTAAAGCAGTCAAGAGACCATGTGTAT 861
QY 300 IleLeuAspArgAsp-----LeuTyrAlaArgSerArgGlyLeu 312
Db 862 GTCAACTCAGTGGACCTCTCGCACAGCGCGCACAGCTTGAGAGCAAGTTCGTTCTCT 921
QY 313 IleSerSerGluGlu 317
Db 922 AAAGCGGCTGCTGAA 936

RESULT 25
US-09-071-035-39
Sequence 39, Application US/09071035
Patent No. 648043
GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:

Db 659 TACGAATATGTTTATAGAAAAATCTCGGATTCCTTTGGTAGATCGCACCAAGCA 718
Qy 272 SerSerAnGluProSerLeuLysGluLeuLysAspProValTrpLysLysLeuAsn 291
Db 719 ATTGGTCGCGAGTATCAAAAGATAACGTCGCTCAAAACGAATTGATTCAAAAACCGAT 778
Qy 292 AlaValLysAnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArgGly 311
Db 779 GCTGTAATAAATGATAAGTCAATATGCTTCAACACGATGTTTGGTATCTAAAGCGGTGT 838
Qy 312 LeuLeSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328
Db 839 GGATTAGAAATCAATGCATTG-----ATGATAGAAGATGTTAAAAA 880
RESULT 26
US-09-602-787A-571
; Sequence 571, Application US/09602787A
; Patent No. 6696561
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Mark
; APPLICANT: Kruger, Burkhard
; APPLICANT: Schder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-125CP
; CURRENT APPLICATION NUMBER: US/09/602,787A
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: USSN 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931454.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931478.0
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931563.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932124.8
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932180.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932182.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932190.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932191.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932212.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932227.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932228.7
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932229.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932927.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14

; PRIOR APPLICATION NUMBER: DE 19940764.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940765.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940766.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940830.0
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940831.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940832.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940833.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941395.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942077.7
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942078.5
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942079.3
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 678
; SEQ ID NO 571
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1018)
; OTHER INFORMATION: RXN03084
US-09-602-787A-571
Alignment Scores:
Pred. No.: 7,48e-21 Length: 1041
Score: 266.00 Matches: 82
Percent Similarity: 44.79% Conservative: 60
Best Local Similarity: 25.87% Mismatches: 135
Query Match: 15.74% Indels: 40
DB: 4 Gaps: 7
US-10-724-972A-6352 (1-335) x US-09-602-787A-571 (1-1041)
Qy 17 LeuPheValLeuIleAlaThrAla-----AlaCysGlyAsnAsnSerSerSer 32
Db 131 CTGATCGTGTCTCGCAGCAGCAGCTACCTCTTACTGCTTGTCTCTCCAGCTCAGAAGAG 190
Qy 33 AsnSerSerLysGluSerSerLysAspGlyValGluIleLysHisGluGluGlyThrThr 52
Db 191 GAAGCATCCACGCTCTGCCACTGCG-----GAATTCACAGACGCTCACGGAACAAC 244
Qy 53 LysValProLysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeu 72
Db 245 GAAGTTCGCGAAATCTCTCGAGCGGTGTTCTCTCGAGCCACTTGAGCTAGACACCGCA 304
Qy 73 ValAlaLeuAspValLysProValGlyIleAlaAspAsnLysLysAsnArgIleIle 92
Db 305 ATCGCCCTCGGAATCACCCAGTGGTGCAGCTGCGCCAAACAGTCCTCCTGGTATT--- 361
Qy 93 LysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsn 112
Db 362 ---CCTGCATATCTCGCGGTGCGATGGAATCAGCTCTCGCGACCTTCTGAGCCAAAT 418
Qy 113 LeuGluGluLeuSerLysLeuLysProAspLeuIleIleAlaAspAsnArgHisLys 132
Db 419 ATCGAAGCGATCGTCTCTCGAGCCGACCTGATCTCTGGCACCAGTATCCCGCACGCC 478

133 GlyIleTyrLysAspLeuAenLysIleAlaProThrIleGluLeuLysSerPheAspGly 152
152
479 GAAATCTACGACCGCTTCGAATCCATCGCCCAACCGTGTTCATGTACAACCCCATGTGAT 538
538
153 AspTyrAsnGluAenLysIleAlaPheLysThrIleSerLysAlaLeuLysGluGlu 172
172
539 CCGTGGAAAGACAATGTC-----GTCCTTCATCGCGCATGTCATGGCCAAAGACGAG 589
589
173 GluGlyLysLysArgLeuGluGluHisAspLysLysIleGluLysLysGluLys 192
192
590 GAATCGGAGATCTCATCCAGCGCTTAATGACAAGTCGCAAGAGATCAAGTCGAGCAT 649
649
193 ThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAla 212
212
650 GATGTCGAAGTAAGACCGCTCAACATGATTCGTCGCCCGCAGCAGCAACCATGAGCCTA 709
709
213 HisProSerAsnSerTyrValGlyGlnPheLysSerGlnLysGlyPheLysGluAlaLeu 232
232
710 TACGGCCGACCTCATTTGGCGGAGCTCTTTGGAGTGGCCAGGA----- 754
754
233 SerAspAspValThrLysGlyLysSerLysTyrLysLysGlyProTyrLeuGlnMetAsn 252
252
755 -----CTCACCATTCTGTATGAGTAATGGAAG 781
781
253 ThrGluThrLeuSerGlnValAsnProGluArgMetPheLe----- 266
266
782 GATGACCTCCAGCGCGACATCGCTCCTGAGAACTTCATGTCTGCCACCACCGCGACTACGTC 841
841
267 ---MetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLysAspPro 285
285
842 TTTGTACCGCAACTGATGTCCACGATGAAATGAGTCTCCCGAAGTAATCCGAGAA--- 898
898
286 ValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeu 305
305
899 -----AACCGGAACAGTTCCTCCATCCTCACCCTTGCTGATACCAGCTAC 943
943
306 TrpAlaArgSerArgGlyLeuLysSerGluGluMetAlaLysGluLeu 322
322
944 TGGGTATCTGGCGTCCGTCCTTGGCGGACGAAAGTCTTTGGAAGACATC 994
994

RESULT 27
US-09-134-001C-2023
; Sequence 2023, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2023
; LENGTH: 183
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2023

Alignment Scores:
Pred. No.: 7,14e-22 Length: 183
Score: 265.00 Matches: 52
Percent Similarity: 96.43% Conservative: 2
Best Local Similarity: 92.86% Mismatches: 2
Query Match: 15.68% Indels: 0
DB: 3 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-134-001C-2023 (1-183)

Qy 280 GluLeuGluLysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAsp 299

Db 23042 CCGATGAGCGGCAAGGACTTCCATCGCTCGGTGATTGCCAACAGGAGATGCTAGCCAGC 22983
 Qy 167 AlaleuGlyLysGluGluGluGlyLysLysArgLeuGluGluHisAspLysLysLeu 186
 Db 22982 GTCCTTGGCAAGGAGGAGGCGCGCGCTGATTGAGGACCTCGCGAAGTCGTGGCG 22923
 Qy 187 GluTyrLysLysGluLeuMetAspLysAsnGlnLysValLeuProAlaValAla 206
 Db 22922 GACCTC---CAACAGATACCGGCACCGCGC---AAGGGCTGGTCTGCTCACC 22869
 Qy 207 LysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeu 226
 Db 22868 GGAGGGCGCATGAGCGCTATGGCGGTTTCGGCTTGGCGTCATCCATCCGTCCTC 22809
 Qy 227 GlyPheLysGluAlaLysAspValThrLysGlyLysLysLysLysLysLysGly 246
 Db 22808 GCGTGGCGGAGGCGCTC-----GAGGGCTCGCCACCTCGCTGCACGGC 22764
 Qy 247 ProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheLe 266
 Db 22763 GAG-----TCCATCAGCGGAGTTCATCCGGAGGAAGATCCCGACTCGCTGCTC 22710
 Qy 267 MetThrAsnLysAlaSerSer-----AsnGluProSerLeuLysGluLeuLysAsp 284
 Db 22709 ATCGACCGGACGCGCCACCGCTCAGAGGAAGGCAATCCCGCAGGTGCTGGACAT 22650
 Qy 285 ProValTrpLysLysLeuAlaValLysAsnGlnArgValAspLeuLeuAspArgAsp 304
 Db 22649 GAATTGGTGGCAGACGACGCGCTGGCAGAAGAACCGAGTCTATCTACCTGGACCCGCGC 22590
 Qy 305 LeuTrpAlaArgSerArgGlyLeuLeuSerSerGluGluMetAlaLysGluLeuVal 323
 Db 22589 GTCACCTACCTGACCGCGCGCATCCAGTCCGTCAGCAGCTCCGCGACCAAGGTC 22533

RESULT 29
 US-09-583-110-1062
 ; Sequence 1062, Application US/09583110
 ; Patent No. 6699703
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al.
 ; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
 ; FILE REFERENCE: Pneumoniae for Diagnostics and Therapeutics
 ; CURRENT APPLICATION NUMBER: US/09/583,110
 ; CURRENT FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/107,433
 ; PRIOR FILING DATE: 1998-06-30
 ; PRIOR APPLICATION NUMBER: US 60/085,131
 ; PRIOR FILING DATE: 1998-05-12
 ; PRIOR APPLICATION NUMBER: US 60/051,553
 ; PRIOR FILING DATE: 1997-07-02
 ; NUMBER OF SEQ ID NOS: 5322
 ; SEQ ID NO 1062
 ; LENGTH: 1035
 ; TYPE: DNA
 ; ORGANISM: Streptococcus pneumoniae
 US-09-583-110-1062

Alignment Scores:
 Pred. No.: 1,77e-19 Length: 1035
 Score: 254.00 Matches: 93
 Percent Similarity: 46.5% Conservative: 56
 Best Local Similarity: 29.0% Mismatches: 119
 Query Match: 15.0% Indels: 52
 DB: 4 Gaps: 14

US-10-724-972A-6352 (1-335) x US-09-583-110-1062 (1-1035)

Qy 13 ValLeGlyLeuPheValLeuLeuAlaThrAlaAlaCysGlyAsnAspSer----- 30
 Db 28 CTAATAGCTATTTAGCTATGATAGTATTTAGCGCTTCTTCTTAATCTGTTAA 87
 Qy 31 ---SerSerAsnSerSerLysGluSerLysAspGlyValGluLeuLysHisGluGlu 49

Db 88 AATGAAGAAATACTTCTAAAGAGCATGCGCTGATATAAATAGTTTATGATCATGCTTTC 147
 Qy 50 GlyThrThrLysValProLysHisProlLysArgValValValLysGluTyrSerPheVal 69
 Db 148 GGTCAAACTATATTAGATAAAAAACCTCAAGAGCTTCAACTATTGCTGGGGAATCAT 207
 Qy 70 AspAlaLeuValAlaLeuAspValLysProValGlyLeuAlaAspAsn----- 86
 Db 208 GATGTACATAGCTTTAGAAATAGTTCCTGTTGGATTTTCAAAAGCAAAATTACGGTGA 267
 Qy 87 LysLysAsnArgLleLysPro-----LeuArgAspLysLysLysLysLys 101
 Db 268 AGTGCTGATAAAGGAGTTTACCACGACAGAGAAAAAATCAAGAACTAAATGGTAAA 327
 Qy 102 TyrThrSerValGlyThrArgLysGlnProAsnLeuGluLysLysLysLysLysPro 121
 Db 328 GCTAACCTATTGACCGATTTGGATGGAATTAACCTTTGAAGCAATATCAAAATCTTAACCA 387
 Qy 122 AspLleLleAlaAspAsnAsnArgHisLysGlyLle-----TyrLysAsp 137
 Db 388 GATGTTATCTTACAGGT-----TATTCGGTATAACTAAAGAAAGATTATGACACT 438
 Qy 138 LeuAsnLysLleAlaProThrLleGluLeuLysSerPheAspGlyAspTyrAsnGluAsn 157
 Db 439 CTATCAAAATTCCTCTGTAGCAGCATCAAAATCTAAACCTTGGCAAACTTTATGGAGA 498
 Qy 158 IleAspAlaPheLysThrLysLysLysAlaLeuGlyLysGluGluGly----- 174
 Db 499 ---GATATGATTTAAATTCATCAAAAGCCTTAGGTATGAAAAAGAGGTGATGAGTTA 555
 Qy 175 -----LysLysArgLysGluGluGluHisAspLysLysLysLysLys 186
 Db 556 ATCAAAAATACTGAAGCTCGTATATCCAAAGAAATTAGAAAAACAT-----CCA 603
 Qy 187 GluTyrLysLysGluLleThrMetAspLysAsnGlnLysValLeuProAlaVal----- 204
 Db 604 GAAATCAAGGAAATC-----AAAGGAAAAAAGATTTATTACTATGATTAAT 654
 Qy 205 ---AlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeu 223
 Db 655 GCTGCAGATACATCAAAATTCGTGATTATATAGCAAGATCCAAAGAGCAAAATTATTTA 714
 Qy 224 SerGlnLeuGly-----PheLysGluAlaLeuSerAspValThrLysGlyLeuSer 241
 Db 715 ACAGATTAGGTCTAGTTTTCCTCGAATCATTTAAAGAAATTTGAGAGTGAA-----GAT 768
 Qy 242 LysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnPro 261
 Db 769 AGTTTTCGAAG-----GAAATTTCTGAGAGAAAGCAAAATAAGATAATGAT 816
 Qy 262 GluArgMetPheLleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu 281
 Db 817 GCTGATGTAATCATA-----ACTTATGCTGATGATATAAACTCTTGAAGCTTTA 864
 Qy 282 GluLysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspLleLeu 301
 Db 865 CAAAGAGATCTCTTTTAGGTAAAAATAATGCAATTAATAAATGTTGCGCTGCTGTAATT 924

RESULT 30
 US-08-961-527-164/c
 ; Sequence 164, Application US/08961527
 ; Patent No. 6420135
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 391
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA

188 Tyr-----LysLysGluIleThrMetAspLysAsnGlnLys 199
6420 GTGCTACTAAAAATGAAAGCTCTGACAAAAAGCCCTTCGCGATCCTCTCTTAATGAAGGA 6361
200 ValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrVal 219
6360 AAAATGGCAGCCTTGGTGCCAAATCTCGTTTC----- 6328
220 GlyGlnPheLeuSerGln---LeuGlyPheLys-----GluAlaLeuSerAspVal 236
6327 ----TCTTTCTTGTACCAACCTTGAATTCACCACTGATACAAATTTGAAGACTCA 6271
237 ThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeu 256
6270 CGCCACGA-----CAAGAAGTCAGCTTTTGAAGTGC 6238
257 SerGlnValAsnProGluArgMetPheIleMetThrAsnLysAla----- 271
6237 AAGAAATCAACCTGATCCTCTTTGTATCAACCGTACCTTCGTCATCGTGGGAC 6178
272 ---SerSerAsnGluProSerLeuLysGlu-----LeuGluLysAspProValTyrLys 288
6177 AACTTAGCAACGAGCGGTCTTAGAAATGCCCTTATCGCTGAAACACCT- 6127
289 LysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTyrAlaArg 308
6126 -----GCTGCTAAAAATGGTAAGATTATCAACTAACACACAGACCTCTGTATCTA 6076
309 SerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328
6075 AGCGAGCGGAGCTTGAATCAACAAACTATGATTGAAGACATACAAAAAGCTTTGAAA 6016
RESULT 31
US-09-583-110-312
; Sequence 312, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 312
; LENGTH: 966
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-312
Alignment Scores:
Pred. No.: 4,6e-19 Length: 966
Score: 250.00 Matches: 92
Percent Similarity: 46.45% Conservative: 65
Best Local Similarity: 27.22% Mismatches: 119
Query Match: 14.79% Indels: 62
DB: 4 Gaps: 14
US-10-724-972A-6352 (1-335) x US-09-583-110-312 (1-966)
QY 20 LeuIleAlaThrAlaAlaCysGlyAsnAsnSerSerSerSerLysGluSerSer 39
DB 49 TTGCTCCTACTTGGTGCATGTAGTACAAACTCAAGCACTAGTCAGACAGACAGCAGTAGC 108
QY 40 LysAspGlyValGlu-----IleLysHisGluGluGlyThrThrLysValProLysHis 57

ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 164:
SEQUENCE CHARACTERISTICS:
LENGTH: 9707 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-164
Alignment Scores:
Pred. No.: 6,67e-18 Length: 9707
Score: 253.00 Matches: 91
Percent Similarity: 46.18% Conservative: 66
Best Local Similarity: 26.76% Mismatches: 117
Query Match: 14.97% Indels: 66
DB: 3 Gaps: 14
US-10-724-972A-6352 (1-335) x US-08-961-527-164 (1-9707)
QY 20 LeuIleAlaThrAlaAlaCysGlyAsnAsnSerSerSerSerLysGluSerSer 39
DB 6930 TTGCTCCTACTTGGTGCATGTAGTACAAACTCAAGCACTAGTCAGACAGACAGCAGTAGC 6871
QY 40 LysAspGlyValGlu-----IleLysHisGluGluGlyThrThrLysValProLysHis 57
DB 6870 TCTGCTCCACAGAGGTAAACCAATTAAGATTCACTGGAGAGGTCAAACTTTCCAAAGTT 6811
QY 58 ProLysArgValValLeuLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspVal 77
DB 6810 CTGAAAGATTGTGACCTTTCACCTTCGCGCTGCGGATCTATTTCGCGCTTTAGGATT 6751
QY 78 LysProValGlyIleAlaAspAspAsnLysLysAsnArgIleIleLysPro----- 94
DB 6750 -----GAAAAAATATCGTCGGAATGCTCAAAAACT 6718
QY 95 -----LeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGln 110
DB 6717 GTTCCGACTTATCTAAAGACCTAGTGGGAACCTGTCAAAAATGTTGGTTCATGAAGAA 6658
QY 111 ProAsnLeuGluGluIleSerLysLysProAspLeuIleIleAlaAspAsnAsnArg 130
DB 6657 CTGATTAGAGACTATCGCGCCCTTGAGCCTGATTGATTATCGCTTCGCGCACGTACA 6598
QY 131 HisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPhe 150
DB 6597 CAAAAATTCGTAGACAAA---TTCAAGAAATCGCCCAACCGTTCTTCTCCAAAGCAAGC 6541
QY 151 AspGlyAspTyrAsnGluAsnIleAspAla---PheLysThrIleSerLysAlaLeuGly 169
DB 6540 AAGGACGACTACTGGGACTTCTTACCAAGGCTTAATATCGAATCCTTAGCAAGTGCCTCGGC 6481
QY 170 Lys-----GluGluGluGlyLysArgLeuGluGluHisAspLysLysIleGluGlu 187
DB 6480 GAAACTGGTACACAGAAAGCCAGGAAGATTGACCAAGCTAGACAAAGCATCCAGAA 6421

Db 109 TCTGCTCCACAGAGGTAACCATTAAGATTCACGTGACGAGGTCAAACTTTCCAAAGTT 168
Qy 58 ProLysArgValValLeuGluTyrSerPheValAlaLeuValAlaLeuAspVal 77
Db 169 CCTGAAAGATTGTGACCTTGGCTCGCGCTCGGATACCTATTCGGCTTTAGGATTT 228
Qy 78 LysProValGlyIleAlaAspAsnLysLysAsnArgIleIleLysPro----- 94
Db 229 -----GAAAAAATATCGTCGGAATGCCTACAAAAACT 261
Qy 95 -----LeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGln 110
Db 262 GTTCCGACTTATCTAAAGACCTAGTCGGAACTGTCAAAATGTGTGTTTATGAAGA 321
Qy 111 ProAsnLeuGluGluIleSerLysLeuLysProAspLeuIleIleAlaAspAsnArg 130
Db 322 CCTGATTAGAGCTATCGCGCTTGGCTGATTTGATTGATTCGTTCCACACGTACA 381
Qy 131 HisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPhe 150
Db 382 CAAAATTCGTAGACAAA---TTCAAGAATAATCGCCCAACCGTTCTTCCCAAGCAAGC 438
Qy 151 AspGlyAspTyrAsnGluAsnIleAspAla---PheLysThrIleSerLysAlaLeuGly 169
Db 439 AAGGACGACTACTGGACTTCTACCAAGGCTAATATCGAATCCTTAGCAAGTCCTTCGGC 498
Qy 170 Lys-----GluGluGluGlyLysArgLeuGluHisAspLysLysIleGluGlu 187
Db 499 GAAACTGGTACACAGAAAGCAAGGAATTCGCCAAGCTAGACAGACATCCAGAA 558
Qy 188 Tyr-----LysLysGluIleThrMetAspLysAsnGlnLys 199
Db 559 GTGCGTACTAAATGAAAGCTCTGCAAAAAGCGCTTCGATCCTCTTAATGAAGA 618
Qy 200 ValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrVal 219
Db 619 AAAATGGCAGCTTTGTCGCAATCTCGTTTC----- 651
Qy 220 GlyGlnPheLeuSerGln---LeuGlyPheLysGluAlaLeuSerAspValThrLys 238
Db 652 ---TCTTCTGTACCAACCTTGAATTCAAACCACT-----GATACAACATTT 699
Qy 239 GlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGln 258
Db 700 GAAGACTCAGCCGACGACAA-----GAAGTCAGCTTTGMAAGTGTCAAGAA 747
Qy 259 ValAsnProGluArgMetPheIleMetThrAsnLysAla-----Ser 272
Db 748 ATCAACCTTGACATCTCTTGTATCAACCGTACCTTGCATCGCTGGTGGGACAACTCT 807
Qy 273 SerAsnGluProSerLeuLysGlu-----LeuGluLysAspProValTrpLysLysLeu 290
Db 808 AGCAACAACGGTGTCTAGAAAATGCCCTTATCGCTGAACACCT----- 852
Qy 291 AsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArg 310
Db 853 ---GCTGTAATAATGGTAAGATTATCAACCTTAACACCCAGACCTCTGGTATCAAGCGA 909
Qy 311 GlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328
Db 910 GCGGACTTGAATCAACAAAATCATGATTGAAGACATACAAAAAGCTTTGAAA 963

RESULT 32
US-09-107-433-2206
; Sequence 2206, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGN
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/ 085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneka
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 2206:
SEQUENCE CHARACTERISTICS:
LENGTH: 978 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (8) LOCATION 1...978
SEQUENCE DESCRIPTION: SEQ ID NO: 2206:
US-09-107-433-2206
Alignment Scores:
Pred. No.: 4,696-19 Length: 978
Score: 250.00 Matches: 92
Percent Similarity: 46.45% Conservative: 65
Best Local Similarity: 27.22% Mismatches: 119
Query Match: 14.79% Indels: 62
DB: 4 Gaps: 14
US-10-724-972A-6352 (1-335) x US-09-107-433-2206 (1-978)
Qy 20 LeuIleAlaThrAlaAlaCysGlyAsnAsnSerSerAsnSerSerLysGluSerSer 39
Db 61 TTGCTCTACTTGTGTCATGTAGTACAAACTCAAGCACTAGTCAGACAGACAGACAGTAGC 120
Qy 40 LysAspGlyValGlu-----IleLysHisGluGluGlyThrThrLysValProLysHis 57
Db 121 TCTGCTCCACAGAGGTAAACCATTAAGTTCACTGGACGAGGTCAAACTTTCCAAAGTT 180
Qy 58 ProLysArgValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspVal 77
Db 181 CCTGAAAGATTGTGACCTTGTACCTCGCGCTCGGATACCTATTCGCGCTTTAGGATTT 240
Qy 78 LysProValGlyIleAlaAspAsnLysLysAsnArgIleIleLysPro----- 94
Db 241 -----GAAAAAATATCGTCGGAATGCCTACAAAAACT 273
Qy 95 -----LeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGln 110
Db 274 GTTCCGACTTATCTAAAGACCTAGTCGGAACCTGCTCAAAATGTGTGTTTATGAAGA 333

[illegible]

Pred. No.: 111e-18 Length: 1119
Score: 247.50 Matches: 95
Percent Similarity: 46.76% Conservative: 71
Best Local Similarity: 26.76% Mismatches: 130
Query Match: 14.64% Indels: 59
DB: 4 Gaps: 16

US-10-724-972A-6352 (1-335) x US-09-602-787A-597 (1-1119)

QY 9 LysileLeu-----SerValileGlyLeuPheValLeuLeuAlaThrAlaAaCys 26
DB 119 AAAATCTGGCAGTAGCGTGGAGGTTGCT---GTCTAGCCCTGTTAGCTGGGTGT 175

QY 27 GlyAsnAsnSer-----SerAsnSerSerLysGluSerLysAspGly 42
DB 176 TCTAAACATGCAGATGACACCGAGCTGATTCAACATCCACCGAAATCCGCTTTTCT 235

QY 43 ValGluileLysHisGluGluGlyThrLysValProLysHisProLysArgValVal 62
DB 236 GTTTCATTAACACAGGTTGGAAACACCAATCGATGATGATCCCGAAAGAGTTGTC 295

QY 63 ValLeuGluTySerPheValAspAlaLeuValAlaLeuAspValLysProValGlyile 82
DB 296 ACCCTTGGCGTTACCGACCGCATATGTCCTCGCATTTGGGACCGTCCAGTAGGCAAC 355

QY 83 AlaAspAsnLysLysAsnArgileileLysProLeuArgAspLysile-----Gly 100
DB 356 ACCGGATACAAATCTCTGAAACAGGATTGGGACCGTGCAGTGTAGTAGTGAAGGC 415

QY 101 Lys---TyrThrSerValGlyThrArgLysGlnProAsnLeuGluLysSerLysLeu 119
DB 416 AAAGAAATTAACACTGCTGCTGACTCTGATTCACACAGATCTTGAACAAGTAGCAGCCCTG 475

QY 120 LysProAspLeuileileAlaAspAsnArgHisLysGly---IleTyLysAspLeu 138
DB 476 GAGCAGACCTGATATTGGAGTCTCTCGGGGTTGACGAGCTGTATACGAGCAACTA 535

QY 139 AsnLysIleAlaProThrileGluLeuLysSerPheAspGlyAspTyraen---GluAen 157
DB 536 TCTGATATCGCACCGGTGTCGCCCGTCCAGCGGAAACAGTCGATACCGAGTAGCTGC 595

QY 158 IleAspAlaPheLysThrileSerLysAlaLeuGlyLysGluGluGlyLysLysArg 177
DB 596 GAGGAGCTACCAACTGTTGTCGCGTGGATGGGGCAATCAGAAAAAGCAAGAGCTC 655

QY 178 LeuGluGluHisAspLysLysileGluGluTyLysLysGluileThrMetAspLysAsn 197
DB 656 AATGAGGAAACAGATGCTGATCCAGCTGCGCGT-----GATGAAAAAT 700

QY 198 GlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisPro-----214
DB 701 -----CCTTCTTTGACGGTAAACACAGAACCGTCTCTGCCATACACAGGCT 748

QY 215 ---SerAsnSerTyrrVal-----GlyGlnPheLeuSerGlnLeuGlyPhe 228
DB 749 AAATACGGTGCTTACCTGCCAGCGATGACCGGGGACAAATTCCTCGATTCACTTGGCATT 808

QY 229 -----LysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyLysLys 245
DB 809 TCGTGCCGGGAGCAGTTCTTTCGGCAGACACCGCGCAGCTTC-----853

QY 246 GlyProTyLysLeuGlnMetAsnThrLysSerGlnValAsnProGluArgMetPhe 265
DB 854 -----TTTGTCGATGTCCTCCGCTGAAAGCGTCAAGACGATAGCGGTGATGTTCTCCTC 907

QY 266 IleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGluAspPro 285
DB 908 GTGCTT-----TCCAAACGACGAAATCTGGATATCACAGCAGAGAAATCCA 952

QY 286 ValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeu 305
DB 953 CTGTTTGAACACTCAACGTTGTGAAAAAGACGCGAGTAATTGTGGCAACACCGAA---1009

QY 306 TrpAlaArgSerArgGlyLeuile-----SerSer 315
DB 1010 -----GAACGGGGGGGATACCTACAACCTAGCTGCTGTCTTGTTCGCTTG 1060

QY 316 GluGluMetAlaLysGluLeuValGluLeuSerLysLysAspSer 330
DB 1061 GAACATCTCGCACCGCTATTGCTGAGGCTTTTGAAGTAAACTCA 1105

RESULT 34
US-09-328-352-318
; Sequence 318, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 318
; LENGTH: 975
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-318

Alignment Scores:
Pred. No.: 5.05e-18 Length: 975
Score: 241.00 Matches: 79
Percent Similarity: 47.42% Conservative: 77
Best Local Similarity: 24.01% Mismatches: 135
Query Match: 14.26% Indels: 38
DB: 4 Gaps: 10

US-10-724-972A-6352 (1-335) x US-09-328-352-318 (1-975)

QY 14 IleGlyLeuLeuPheValLeuileAlaThrAlaAaCysGlyAsnAsnSerSerSerAsn 33
DB 34 GTCGCCCTCATCAGCGCGCTCTCACTTACAGCATGTGATCAAAAGTTGCAGAT 93

QY 34 SerSerLysGluSerSerLys-----AspGlyValGluileLysHisGluGlyThr 51
DB 94 ACCACTCAAGCTTCTCAGAGCTAGCAGACCGATTACGGTAAACATCGCGTTGGAAT 153

QY 52 ThrLysValProLysHisProLysArgValValValLeuGluTySerPheValAspAla 71
DB 154 ACGGTGATAGATCATCTACTACCGCAACCGCTTGCCTGCTTATGATGATGAATGAAGCCGACTTT 213

QY 72 LeuValAlaLeuAspValLysProValGlyileAlaAspAspAsnLysLysAsnArgile 91
DB 214 CTAGATCAACTCAATGTCCCGATTATGGGAATGCCAAAGATTATGTTCCACACTTTTG 273

QY 92 IleLysProLeuArgAspLysileGlyLysTyThrSerValGlyThrArgLysGlnPro 111
DB 274 GAAAAATATAAAAGAT-----GCACAAATTCAGGATTTGGGTGCAATTGTACAGCCC 327

QY 112 AsnLeuGluGluileSerLysLysLeuLysProAspLeuileleAlaAspAsnAsnArgHis 131
DB 328 AATATGAAAGGATATATGATTAACACACACCTGATTTTGATG---ACGCCATTACAC 384

QY 132 LysGlyileTyLysAspLeuAsnLysileAlaProThrileGluLeuLysSerPheAsp 151
DB 385 GTTAATCAGTATCAGGAACCTGTCAAAATTTGCCCGCACCATTCAT-----TACGAT 435

QY 152 GlyAspTyraenGluAsn-----IleAspAlaPheLysThr 163
DB 436 ATTAACCTTCAACATACCGAGAGTAAATCATATTGCTGTGTTAAAGACCATATGATGACC 495

QY 164 IleSerLysAlaLeuGlyLysGluGluGlyLysArgLeuGluGluHisAspLys 183
DB 496 TTAGAAAAATATTATTAATAAAGAGATTAGCCCGCCGCAAGAGTTCTGAACTTGATGAA 555

QY 184 LysIleGluGluTyLysLysGluileThrMetAspLysAsnGlnLysValLeuProAla 203

Alignment Scores: 6.36e-18 Length: 954
Pred. No.: 240.00 Matches: 99
Score: 45.28% Conservative: 64
Percent Similarity: 27.50% Mismatches: 103
Best Local Similarity: 14.20% Indels: 94
Query Match: 3 Gaps: 20
DB:

US-10-724-972A-6352 (1-335) x US-09-071-035-69 (1-954)

QY 8 LeuLysIleLeuSerValIleGly-----LeuLeuPheValLeuIleAlaThrAlaAala 25
DB 25 TTAATAAAGACGGTCTCTAATTTGGTACACCCCTTCTTCTGGTTCATTCTTACTCGCAGCT 84
QY 26 CysGlyAasn---AasnSerSerAasnSerSerLysGluSerSerLysAspGlyValGlu 44
DB 85 TGTGGTAATACGTAATAAGAACCAACACGCTGACAAACA-----LysValProLysHisProLys 59
QY 45 IleLysHisGluGluGlyThr-----LysValProLysHisProLys 59
DB 127 -----CATGAAGTAAACAGATACCTTAGGCAATAAAGTAAACCGTCCCGGAAACCCAAA 180
QY 60 ArgValValVal-----LeuGluTyrSerPheValAspAlaLeuValAlaLeuAspVal 77
DB 181 CGGATTAATTCGGAGTTATTAGAA-----GATTATCTAGTTGCATTAGGAGAA 228
QY 78 LysProValGly-----IleAlaAspAspAasnLysLysAasnArgIleIleLysPro 94
DB 229 AAACCACTGGCACAATGGACAGTTGGACAAGGACGATTCAGGATTTATTAGCGAAAGAA 288
QY 95 LeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAasnLeu--- 113
DB 289 TTGAAGAT-----GTCCCCACCTATTTC 312
QY 114 -----GluGluIleSerLysLysProAspLeuIleIleAlaAasn 128
DB 313 TATGACTTGCATATGAGCGGCTTCTAAATTTGAACCTGACTTATTATTATCAATCAATCA 372
QY 129 AasnArg-----HisLysGlyIleTyrLysAspLeuAasnLysIleAlaProThrIleGlu 146
DB 373 TCTGCTCTAGTTGAAGCGGTAATAACAAAGATACAGTAAATAATTCGCCCACTTATGTA 432
QY 147 LeuLysSerPheAspGlyAspTyrAasnGluAasnIle-----AspAlaPheLysThr 163
DB 433 GTCAAAAAC-----GGCGAAAATGTCACTGGCGGTGATCAATTTGGAGAT 477
QY 164 IleSerLysAlaLeuGlyLysGluGlyLysLysArgLeuGluGluHisAsp--- 182
DB 478 ATTGCCACTGTTTATAGATAAAGAACCAAGCAAGGAAAGTCTTAGAAGATTATGATACC 537
QY 183 -----LysLysIleGluGluTyr-----LysLysGluIleThrMetAspLysAasnGln 198
DB 538 TTAACCAAGCGGTCCAGAAATATCTTGGCAAAAAGAT----- 576
QY 199 LysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAasnSerTyr 218
DB 577 -----GCTGGCAATCTCGCGCAGTCTTATGGTAAACCAACCAACCAA 618
QY 219 Val-----GlyGlnPheLeuSerGln---LeuGlyPhe 228
DB 619 GTCTTTATGGTTAGCGATTAATCGCTCAAGCGAACCGTCTCTATCAGGACTTAGGCCTC 678
QY 229 LysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLysLysGlyProTyr 248
DB 679 CAA-----GTTCCAAAATTAGTGGAAAGAAATTTCTAAACCAACCTACTCGCGGATTGG 729
QY 249 LeuGlnMetAasnThrGluThrLeuSerGlnValAasnProGluArgMetPheIleMetThr 268
DB 730 AATCAAGTTTCTTTAGAAAAATTAGCTGAGCTTGGACGACACACATTTTCTTGTGAAC 769
QY 269 AasnLysAlaSerSerAasnGluProSerLeuLysGluLeuLysAspProValTrrLys 288
DB 790 AGCGATGAATCAGCA-----CCTCTTTTCCAGAAAGCAATTTGGAAG 831

DB 556 CAAGTGAAGCAAGTA---CAAGCGGTAAACGGCCCAATCGCCACCAAGAGCATTAGTCGTT 612
QY 204 ValAlaAlaLysSerGlyLeuLeuAlaHisProSerAasnSerTyrValGlyGlnPheLeu 223
DB 613 CTTTCAACAATCGGGCATTAGTAAATTTGGTATTTCAGTCAGCTATGCTTTATTTT 672
QY 224 SerGlnLeuGlyPheLysGluAla-----LeuSerAspAspValThrLysGlyLeuSer 241
DB 673 AATGCCTTTGGGTAAAGCCTGCAAGTGGTGGTGGATACAAAGCTACATGGG----- 726
QY 242 LysTyrLeuLysGlyProTyrLeuGlnMetAasnThrGluThrLeuSerGlnValAasnPro 261
DB 727 -----CAACCTATTTTCGAGTGAATTCATCAAAAAGCGGATCCC 765
QY 262 GluArgMetPheIle-----MetThrAasnLysAlaSerSerAasnGlu 275
DB 766 GATATCTTTATATCGTGGACCGAAGTCTGTGATGGAACATCGCTCAAAACATTAATGCT 825
QY 276 ProSerLeuLysGluLeuLysAspProValTrrLysLysLysLeuAasnAlaValLysAasn 295
DB 826 GCCAGTGT-----GAAAATCCATTTATTACGTCAAAACCAAGCATGGAAAGAAC 873
QY 296 GlnArgValAspIleLeuAspArgAspLeuTrrAlaArgSerArgGlyLeuIleSerSer 315
DB 874 GGTCTGTGTCAATTTTGTGATGCGCGATGCTGCTACACCGGCTGCAAGCCCAACCTCA 933
QY 316 GluGlnMetAlaLysGluLeuValGlu 324
DB 934 CTCAAATAGTGTGAAGATGTGAAA 960

RESULT 35
US-09-071-035-69
; Sequence 69, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brooks
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 954 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-071-035-69

Qy 289 LysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArg 308
Db 832 AACTTACTGCTGTGAAATAAACAAGTTTATACCTATGATAA-----AAA 879
Qy 309 SerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerIysLys 328
Db 880 AGTAGTTGTTATACACAGCGACTATTCCGAATACTCAATTTGTTGAAGATGTAATAAAA 939

RESULT 36

US-09-902-540-5089
; Sequence 5089, Application US/09902540
; Patent No. 6833447

GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B

; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/217,883

; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 16825

; SEQ ID NO 5089

; LENGTH: 837

; TYPE: DNA

; ORGANISM: Myxococcus xanthus

US-09-902-540-5089

Alignment Scores:

Pred. No.: 8,97e-18 Length: 837
Score: 238.00 Matches: 76
Percent Similarity: 49.65% Conservative: 64
Best Local Similarity: 26.95% Mismatches: 126
Query Match: 14.08% Indels: 16
DB: 4 Gaps: 9

US-10-724-972A-6352 (1-335) x US-09-902-540-5089 (1-837)

Qy 45 IleLysHisGluGluGlyThrLysValProLysHisProLysArgValValLeu 64
Db 4 GTGACCGCAGGACAGGACACCGCTGCTCCGCTGAACCGCGGGTGTGTTC 63
Qy 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84
Db 64 GACCTGTGTGGCTGGACATCTCCAGGCCCTGGAGTGGAGCTCCAGCGGTGGCGGC 123
Qy 85 AspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSer 104
Db 124 GACATGTTCCCGCCAGCACCTGACCCGG---TTCCGGGAC-----GCCAAGTACCGCGC 174
Qy 105 ValGlyThrArgLysGlnProAsnLeuGluIleSerLysLeuLysProAspLeuIle 124
Db 175 ATGGCACCTGTTTCGAGCCTGACTACGAGGCACCTCCAGCGCGCGAAGCGGACCTCATC 234
Qy 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144
Db 235 ATCACC---GGGGCGCGTCCAGCGCGAAGTACTCGAACCTTCTCCGCATCGCGCCACC 291
Qy 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAla---PheLysThr 163
Db 292 ATCCAGCTGCGGATGAGCGGAAGGACCTTCATCGCTCGGTGATTGCCAACCGAGATG 351
Qy 164 IleSerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAspLys 183
Db 352 CTAGCCAGCTGTTTGGCAAGGAGGAGGCGCGCGCTGATTGAGGACCTCGCGAAG 411
Qy 184 LysIleGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 203
Db 412 TCGGTGGCGGACCTC---CAACAGATACCGCGCACCGCGC---AAGGGGCTGTGCTG 465

Qy 204 ValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeu 223
Db 466 CTCGTACCGAGGCGCATGAGCGCTATATGCGCGGTTTGGCGTTCATCCAT 525
Qy 224 SerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyr 243
Db 526 GGTGACTTCGGCGTGGCCGAGCGCTC-----GAGGGCTCCGACCTCG 570
Qy 244 LeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263
Db 571 CTGCACGCGAG-----TCCATCAGCGCGAGTTTCATCCGGGAGAGAAGATCCCGACTGG 624
Qy 264 MetPheIleMetThrAsnLysAlaSerSer-----AsnGluProSerLeuLysGluLeu 281
Db 625 CTGTTGCTATCAGCGGACCGCGCCCGTCCAGAAAGGCAATGCCCGCAGGTG 684
Qy 282 GluLysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeu 301
Db 685 CTGGACAATGAATTGGTGGCGGACAGCGCTCCGAGAGAACAGGTCATCTACTCTG 744
Qy 302 AspArgAspLeuTrpAlaArgSerArgLysLeuIleSerSerGluGluMetAlaLysGlu 321
Db 745 GACCCCGCGTCTACCTACCTGACCGCGCGGCGATCCAGTCCGTCAGAGAGTCCCGCGAC 804
Qy 322 LeuVal 323
Db 805 CAGGTC 810

RESULT 37

US-08-961-083-23

; Sequence 23, Application US/08961083

; Patent No. 6159469

; GENERAL INFORMATION:

; APPLICANT: Choi et. al.

; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

; NUMBER OF SEQUENCES: 452

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/961,083

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Brookes, A. Anders

; REGISTRATION NUMBER: 36,373

; REFERENCE/DOCKET NUMBER: PB340P2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 23:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 895 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

US-08-961-083-23

Alignment Scores:

Pred. No.: 1.66e-17 Length: 895

Score: 236.00 Matches: 88

Percent Similarity:	46.22%	Conservative:	65
Best Local Similarity:	26.59%	Mismatches:	112
Query Match:	13.96%	Indels:	66
DB:	3	Gaps:	14
US-10-724-972A-6352 (1-335) x US-08-961-083-23 (1-895)			
QY	29	AsnSerSerSerAsnSerSerLysGluSerSerLysAspGlyValGlu-----lleLys	46
DB	8	AACTCAAGCACTAGTCAGACAGAGACGAGTAGCTCTGCTCCACAGAGGTAAACATATAA	67
QY	47	HisGluGluGlyThrThrLysValProLysHisProLysArgValValValLysGluTyr	66
DB	68	AGTTCACTCGACGAGGTCAAACTTTCAAAGTCTCTGAAAGATTGTGACCTTTGACCTC	127
QY	67	SerPheValAspAlaLeuAlaLeuAspValLysProValGlyLysAlaAspAsn	86
DB	128	GGCGCTCGGATACTATTCGGCTTTAGGATT-----	160
QY	87	LysLysAsnArgLleLysPro-----leuArgAspLysIle	99
DB	161	GAATAAATAATCGTCGGAATGCTTACAAAACTGTTCCGACTTATCTAAAAGACCTAGTG	220
QY	100	GlyLysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluLysSerLysLeu	119
DB	221	GGAACTGTCAAAAATGTTGTTCTATGAAGAACTGATTAGAGCTATCGCGCCCTT	280
QY	120	LysProAspLysLleLysAlaAspAsnArgHisGlyLysLysLysLysLysLysLys	139
DB	281	GAGCTGATTGATTATTCGCTTCGCCAGCTACACAAAAATTCGTAGACAAA---TTCAAA	337
QY	140	LysLysAlaProThrIleGluLysSerPheAspGlyAspTyrAsnGluAsnIleAsp	159
DB	338	GAATCGCCCAACCGTTCTCTTCAAGCAAGCAGCAGCTACTGGAATCTTACCAAG	397
QY	160	Ala---PheLysThrIleSerLysAlaLeuGlyLys-----GluGluGluGlyLysLys	176
DB	398	GCTAATATCGAATCTCTAGCAAGTTCCTCGCGGAACTGTTACACAGAAAGCCAGAA	457
QY	177	ArgLeuGluGluHisAspLysLysLysLysLysLysLysLysLysLysLysLysLys	188
DB	458	GAATGACCAAGCTAGACAGAGCATCCAGAGTTCGCTACTTAAAAATGAAGCTCTGAC	517
QY	189	LysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSer	208
DB	518	AAAAAAGCCCTTCGATCTCTCTTAATGAAGGAAAAATGGCAGCCCTTGGTCCCAATCT	577
QY	209	GlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGln---LeuGly	227
DB	578	CGTTTC-----TCTTCTTGTACCAAACTTGAAA	607
QY	228	PheLys-----GluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLys	245
DB	608	TTCAAAACCACTGATACAAAATTTGAGACTCAGCCACGGA-----	649
QY	246	GlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPhe	265
DB	650	-----CAAGAACTCAGCTTTGAAAGTGTCAAAAGATCAAACTCAGCTCCTCTTT	700
QY	266	IleMetThrAsnLysAla-----SerSerAsnGluProSerLeuLys	279
DB	701	GTCAATCAACCGTACCTTCGATCGTGGGACAACTCTAGCAACAGCGGTGCTTAGAA	760
QY	280	Glu-----LeuGluLysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArg	297
DB	761	AATGCCCTTATCCTGAAACCT-----GCTGTAATAAATGGTAAG	802
QY	298	ValAspIleLeuAspArgAspLeuTrpAlaArgSerArgGlyLeuLysSerSerGluGlu	317
DB	803	ATTATCAACTAACACCAAGCTCTGCTATCTAAGCGGAGCGGACTTGAATCAACAAA	862
QY	318	MetAlaLysGluLeuValGluLeuSerLysLys	328

DB	863	CTCATGATTGAAGACATACAAAAAGCTTTGAAA	895
RESULT 38			
US-09-536-784-23			
; Sequence 23, Application US/09536784			
; Patent No. 6573082			
; GENERAL INFORMATION:			
; APPLICANT: Choi et. al.			
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines			
; NUMBER OF SEQUENCES: 452			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Human Genome Sciences, Inc.			
; STREET: 9410 Key West Avenue			
; CITY: Rockville			
; STATE: Maryland			
; COUNTRY: USA			
; ZIP: 20850			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage			
; COMPUTER: HP Vectra 486/33			
; OPERATING SYSTEM: MSDOS version 6.2			
; SOFTWARE: ASCII Text			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/09/536,784			
; FILING DATE: 30-Oct-1997			
; CLASSIFICATION: <Unknown>			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: 08/961,083			
; FILING DATE: OCT-30-1997			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Michelle S. Marks			
; REGISTRATION NUMBER: 41,971			
; REFERENCE/DOCKET NUMBER: PB340P3			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (301) 309-8504			
; TELEFAX: (301) 309-8512			
; INFORMATION FOR SEQ ID NO: 23:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 895 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: double			
; TOPOLOGY: linear			
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:			
US-09-536-784-23			
Alignment Scores:			
Pred. No.:	1,66e-17	Length:	895
Score:	236.00	Matches:	88
Percent Similarity:	46.22%	Conservative:	65
Best Local Similarity:	26.59%	Mismatches:	112
Query Match:	13.96%	Indels:	66
DB:	4	Gaps:	14
US-10-724-972A-6352 (1-335) x US-09-536-784-23 (1-895)			
QY	29	AsnSerSerSerAsnSerSerLysGluSerSerLysAspGlyValGlu-----lleLys	46
DB	8	AACTCAAGCACTAGTCAGACAGAGACGAGTAGCTCTGCTCCACAGAGGTAAACATATAA	67
QY	47	HisGluGluGlyThrThrLysValProLysHisProLysArgValValValLysGluTyr	66
DB	68	AGTTCACTCGACGAGGTCAAACTTTCAAAGTTCCTGAAAAAGATTGTGACCTTTGACCTC	127
QY	67	SerPheValAspAlaLeuValAlaLeuAspValLysProValGlyLysAlaAspAsn	86
DB	128	GGCGCTCGGATACTATTCGGCTTTAGGATT-----	160
QY	87	LysLysAsnArgLleLysPro-----leuArgAspLysIle	99
DB	161	GAATAAATAATCGTCGGAATGCTTCAAAAGTCTTATCTAAAAGACCTAGTG	220
QY	100	GlyLysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluLysSerLysLeu	119

Db 221 GGAACGTGCAAAATGTTGGTTCTATGAAAGACCTGATTAGAAAGCTATCGCCGCCCTT 280
Qy 120 LysProAspLeuIleIleAlaAspAsnArgHisLysGlyIleTyrLysAspLeuAsn 139
Db 281 GAGCTGATTGATTATCGCTTCGCCACGTACACAAAATTCGTAGACAAA--TTCAA 337
Qy 140 LysIleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAsp 159
Db 338 GAAATGCCCAACCGTCTCTTCCAGCAAGCAAGCACTACTGGACTTCTTACCAAG 397
Qy 160 Ala---PheLysThrIleSerLysAlaLeuGlyLys-----GluGluGluGlyLysLys 176
Db 398 GCTAATATCAATCTTTCAGCAAGTCTTTCGCGCAAACTGGTACACAGAAAGCCAGGAA 457
Qy 177 ArgLeuGluGluHisAspLysLysIleGluGluTyr----- 188
Db 458 GAATTGACCAAGCTAGACAGAGCATCCAAAGAACTCGCTACTAAAAATGAAAGCTCTGAC 517
Qy 189 LysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSer 208
Db 518 AAAAAGCCCTTCGATCTCTCTTAATGAAGGAAAAATGCGACCCCTTGGTGCCAAATCT 577
Qy 209 GlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGln---LeuGly 227
Db 578 CGTTTC-----TCCTTCTGTACCAAACTTGA 607
Qy 228 PheLys-----GluAlaLeuSerAspAspValThrLysGlyLeuSerLysTyrLeuLys 245
Db 608 TTCAAACCAACTGATACAAAATTTGAAGACTCACGCCACGGA----- 649
Qy 246 GlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPhe 265
Db 650 -----CAGAGTCAGCTTTGAAAGTGTCAAAAGAAATCAACCCCTGACATCTCTTT 700
Qy 266 IleMetThrAsnLysAla-----SerSerAsnGluProSerLeuLys 279
Db 701 GTCATCAACCGTACCTCCATCGTGGGCAACTCTAGCAACACCGGTCTCTAGAA 760
Qy 280 Glu-----LeuGluLysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArg 297
Db 761 AATGCCCTTATCGCTGAACACCT-----GCTGCTAAAAATCGTAAAG 802
Qy 298 ValAspIleLeuAspArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGlu 317
Db 803 ATTATCCAACCTAACACCGACCTCTGGTATCTTAAGCGGAGCGGACTTGAATCAACAAA 862
Qy 318 MetAlaLysGluLeuValGluLeuSerLysLys 328
Db 863 CTCATGATTGAAGACATACAAAAGCTTTGAAA 895

RESULT 39

US-09-134-000C-2609
; Sequence 2609, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2609
; LENGTH: 906
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-2609

Alignment Scores:

Pred. No.: 1.7e-17 Length: 906

Score: 236.00 Matches: 93
Percent Similarity: 45.61% Conservative: 63
Best Local Similarity: 27.19% Mismatches: 94
Query Match: 13.96% Indels: 92
DB: 4 Gaps: 19

US-10-724-972A-6352 (1-335) x US-09-134-000C-2609 (1-906)

Qy 24 AlaAlaCysGlyAsn---AsnSerSerSerAsnSerSerLysGluSerSerLysAspGly 42
Db 31 GCGCGCTGTGGTAAATACGAATAAAGAACCAACCAACCTGCACAAACA----- 78
Qy 43 ValGluIleLysHisGluGluGlyThrThr-----LysValProLysHis 57
Db 79 -----CATGAAGTAAACAGATACCTTAGGCAATAAAGTAAACCGTCCCGCGAAA 126
Qy 58 ProLysArgValValVal-----LeuGluTyrSerPheValAspAlaLeuValAlaLeu 75
Db 127 CCCAAACCGATTATTTCGAGTTATTAGAA-----GATTATCTAGTTGCATTA 174
Qy 76 AspValLysProValGly-----IleAlaAspAsnLysLysAsnArgIleIle 92
Db 175 GGAGAAAAAACCAAGTGGCAATGGACAGTTGGACAAGCGCATTTCAAGATTATTAGCG 234
Qy 93 LysProLysArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsn 112
Db 235 AAGAAATTGAAAGAT-----GTCCCCACT 258
Qy 113 Leu-----GluGluIleSerLysLeuLysProAspLeuIleIleAla 126
Db 259 ATTCTCTATGACTTGCATATGAGCGGTTCTAAATTTGAACCTGACTTATTATTATC 318
Qy 127 AspAsnAsnArg-----HisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144
Db 319 AGTTATCTGCTCTAGTTGAAGCGGTAAATACAAAGATAACAGTAAATTTGCGCCACT 378
Qy 145 IleGluLeuLysSerPheAspLysAspTyrAsnGluAsnIle-----AspAlaPhe 161
Db 379 TATGTAGTCAAAAAC-----GCCGAAAATGTCACTGCGGTGATCAATTG 423
Qy 162 LysThrIleSerLysAlaLeuGlyLysGluGluGlyLysLysLysArgLeuGluHis 181
Db 424 GAAGATATTGCCACTGTTTATAGATAAAGAAAGCAAGCGAANAAGTGTAGAAAGATTAT 483
Qy 182 Asp-----LysLysIleGluGluTyr-----LysLysGluIleThrMetAspLys 196
Db 484 GATACCTTTAAACCAAGCGCTCAAGAATATCTTGCAAAAAGAT----- 528
Qy 197 AsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsn 216
Db 529 -----GCTGGCAAAATCTGCGCAGTCTTATGGTAAACCAAC 564
Qy 217 SerTyrVal-----GlyGlnPheLeuSerGln---Leu 226
Db 565 AACCAAGTCTTTATGTTAGCGATAATCGCTCAAGCGGAACCGTGCTCTATCAGGACTTA 624
Qy 227 GlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGly 246
Db 625 GGCCTCCAA-----GTTCCAAAATTAGTGAAGAATTTCTAAACCGCTACTCGC 675
Qy 247 ProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheIle 266
Db 676 GATTGGAATCAAGTTTCTTTAGAAAAATTAGCTGAGCTTGACGACACCATTTTCTCT 735
Qy 267 MetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLysAspProVal 286
Db 736 GTAAACAGCGATGAATCAGCA-----CCTCTTTTCCAAAGAACCAATT 777
Qy 287 TrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrp 306
Db 778 TGAAGAAGAACTTACCTGCTGTGAAAAAATAACCAAGTTCATACCTATGATAAA----- 828
Qy 307 AlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSer 326

146 AAACAGTGGGCAATGACAGTGGACAGGCGACGATTTCAAGATTATTTAGCGAAAGAA 205
95 LeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsnLeu--- 113
206 TTGAAAGAT-----GTCCCACTATTTC 229
114 -----GluGluIleSerLysLeuLysProAspLysIleIleAlaAspAsn 128
230 TATGACTTGCATATGACGCGTTCTAAATTTGAACCTGACTTATATATATATCACTTCA 289
129 AsnArg-----HisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGlu 146
290 TCTGCTCTAGTGAAGCGGTAAATACAAAGATACAGTAAATTCGCCCACTTATGTA 349
147 LeuLysSerPheAspGlyAspTyrAsnGluAsnIle-----AspAlaPheLysThr 163
350 GTCAAAAAC-----GCCGAAATGTCCCTGGCGTGTCAATTTGGAAGAT 394
164 IleSerLysAlaLeuGlyLysGluGluGlyLysArgLysGluGluHisAsp--- 182
395 ATTGCCACTGTTTTAGATATAAAGAACCAAGCGCAAAAGTGTAGAAAGATTATGATACC 454
183 -----LysLysIleGluGluTyr-----LysLysGluIleThrMetAspLysAsnGln 198
455 TTAACCAAGCGCTCCAGGAATATCTTGGCAAAAAGAT----- 493
199 LysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyr 218
494 -----GCTGGCAATCTCGCGCAGTCTTATGGGTAAACCAACAA 535
219 Val-----GlyGlnPheLeuSerGln---LeuGlyPhe 228
536 GTCTTTATGTTAGCGATAATCGCTCAAGCGGAACCGTCTCTATCAGGACTTAGGCTC 595
229 LysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyr 248
596 CAA-----GTTCCAAAATTTAGTGAAGAAATTTCTAAAAACGCTACTGCGATTGG 646
249 LeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThr 268
647 AATCAAGTTTCTTTAGAAAAAATTAGCTGAGCTTGACGACGACACATTTCTTGTAAAC 706
269 AsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLysAspProValTyrLys 288
707 AGCGATGAATCAGCA-----CCTCTTTTCCAGAAAGCAATTTGGAAG 748
289 LysLeuAsnAlaValLysAsnGluArgValAspIleLeuAspArgAspLeuTrpAlaArg 308
749 AACTTACCTGCTGTGAAAAAATAACCAAGTTTCATCTATGATAA-----AAA 796
309 SerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328
797 AGTAGTTGTTATACACCGACCTATTGCGAATACTCAAAATTTGTAAGATGTAAAAAAA 856

Search completed: November 9, 2005, 19:57:49
Job time : 252 secs

Db 829 ---AAAAGTAGTGGTTATACACGCGACCTATTGCAATCACTCAAAATTTGAGAGATGTA 885
Qy 327 LysLys 328
Db 886 AAAAAA 891
RESULT 40
US-09-071-035-71
Sequence 71, Application US/09071035
Patent No. 6448043
GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071.035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 868 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-071-035-71
Alignment Scores:
Pred. No.: 1.32e-16 Length: 868
Score: 228.00 Matches: 91
Percent Similarity: 45.29% Conservative: 63
Best Local Similarity: 26.76% Mismatches: 94
Query Match: 13.49% Indels: 92
DB: 3 Gaps: 19
US-10-724-972A-6352 (1-335) x US-09-071-035-71 (1-868)
Qy 26 CysGlyAsn---AsnSerSerAsnSerLysGluSerLysAspGlyValGlu 44
Db 2 TGTGTAATACGAATAAAGAACCAACGCTGACAAAAACA----- 43
Qy 45 IleLysHisGluGluGlyThr-----LysValProLysHisProLys 59
Db 44 -----CATGAAGTACAGTACCTTAGGCAATAAAGTAAACCGTCCCGCGAAACCCAAA 97
Qy 60 ArgValValVal-----LeuGluTyrSerPheValAspAlaLeuValAspVal 77
Db 98 CGGATTATTGCGAGTTATTAGAA-----GATTATCTAGTTGATTAGGAGAA 145
Qy 78 LysProValGly-----IleAlaAspAsnLysLysAsnArgIleIleLysPro 94

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 10, 2005, 00:11:38 ; Search time 215 Seconds
(without alignments)
2549.547 Million cell updates/sec

Title: US-10-724-972A-6352
Perfect score: 335
Sequence: 1 GVESVRLKILSVIGLFLVL.....EEMAKELVSLKSKKONK 335

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2400006

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Command line parameters: -DEV=xlh

-MODE=frame+pn.model -DEV=xlh
-Q=/cgn2_1/USPTO spool_h/US10724972/runat_03112005_141928_29427/app_query.fasta_1.519
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=oligo rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=1000 -DOALIGN=200 -THR SCORE=quality -THR_MIN=1 -ALIGN=40 -MODE=LOCAL
-OUTFWT=PCO -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10724972@cgn 1.1.69 @runat_03112005_141928_29427 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	15.2	183	3	US-09-134-001C-2023
2	35	10.4	3046	4	US-09-710-279-3782
3	26	7.8	242	4	US-08-956-171E-2556
4	26	7.8	242	4	US-08-781-986A-2556
5	23	6.9	668	4	US-08-956-171E-355
6	23	6.9	668	4	US-08-781-986A-355
7	14	4.2	999	4	US-09-830-217-15
8	14	4.2	999	4	US-10-278-946-15
9	14	4.2	3775	4	US-08-956-171E-238
10	14	4.2	3775	4	US-08-781-986A-238
C 11	9	2.7	13248	4	US-09-949-016-12024
C 12	9	2.7	13248	4	US-09-949-016-14230

C 13	9	2.7	203475	4	US-09-949-016-14516	Sequence 14516, A
C 14	9	2.7	203475	4	US-09-949-016-14517	Sequence 14517, A
C 15	9	2.7	203475	4	US-09-949-016-14518	Sequence 14518, A
C 16	9	2.7	203475	4	US-09-949-016-14519	Sequence 14519, A
C 17	9	2.7	203475	4	US-09-949-016-17226	Sequence 17226, A
C 18	9	2.7	203475	4	US-09-949-016-17227	Sequence 17227, A
C 19	9	2.7	203475	4	US-09-949-016-17228	Sequence 17228, A
C 20	9	2.7	203475	4	US-09-949-016-17229	Sequence 17229, A
C 21	8	2.4	181	4	US-09-513-999C-3507	Sequence 3507, A
C 22	8	2.4	305	4	US-09-513-294A-5824	Sequence 5824, Ap
C 23	8	2.4	352	4	US-09-513-999C-2950	Sequence 2950, Ap
C 24	8	2.4	378	4	US-09-513-999C-29145	Sequence 29145, A
C 25	8	2.4	390	4	US-09-328-352-884	Sequence 884, App
C 26	8	2.4	564	4	US-09-902-540-2334	Sequence 2334, Ap
C 27	8	2.4	601	4	US-09-949-016-39108	Sequence 39108, A
C 28	8	2.4	601	4	US-09-949-016-60203	Sequence 60203, A
C 29	8	2.4	601	4	US-09-949-016-60204	Sequence 60204, A
C 30	8	2.4	601	4	US-09-949-016-135038	Sequence 135038, A
C 31	8	2.4	601	4	US-09-949-016-149234	Sequence 149234, A
C 32	8	2.4	789	3	US-09-134-001C-909	Sequence 909, App
C 33	8	2.4	789	4	US-09-710-279-1567	Sequence 1567, Ap
C 34	8	2.4	843	4	US-09-543-681A-1850	Sequence 1850, Ap
C 35	8	2.4	886	4	US-09-270-767-14140	Sequence 14140, A
C 36	8	2.4	981	4	US-09-489-039A-2246	Sequence 2246, Ap
C 37	8	2.4	1001	3	US-09-641-638-192	Sequence 192, App
C 38	8	2.4	1001	3	US-09-641-638-193	Sequence 193, App
C 39	8	2.4	1001	4	US-10-170-097-192	Sequence 192, App
C 40	8	2.4	1001	4	US-10-170-097-193	Sequence 193, App
C 41	8	2.4	1389	4	US-09-489-039A-3577	Sequence 3577, Ap
C 42	8	2.4	1470	4	US-09-902-540-133	Sequence 133, App
C 43	8	2.4	1491	4	US-09-134-000C-497	Sequence 497, App
C 44	8	2.4	1875	3	US-09-422-869-21	Sequence 21, Appl
C 45	8	2.4	2064	4	US-09-620-312D-375	Sequence 375, App
C 46	8	2.4	2103	4	US-09-620-312D-376	Sequence 376, App
C 47	8	2.4	2148	4	US-09-620-312D-374	Sequence 374, App
C 48	8	2.4	2283	4	US-09-248-796A-3465	Sequence 3465, Ap
C 49	8	2.4	3047	4	US-09-710-279-4113	Sequence 4113, Ap
C 50	8	2.4	3100	4	US-09-270-767-12855	Sequence 12855, A
C 51	8	2.4	3572	4	US-09-270-767-11151	Sequence 11151, A
C 52	8	2.4	3821	4	US-08-956-171E-304	Sequence 304, App
C 53	8	2.4	3821	4	US-08-781-986A-304	Sequence 304, App
C 54	8	2.4	5648	3	US-09-371-008-1	Sequence 1, Appl
C 55	8	2.4	13015	4	US-09-639-207-10	Sequence 10, Appl
C 56	8	2.4	17607	4	US-09-949-016-15968	Sequence 15968, A
C 57	8	2.4	35524	3	US-08-923-137-1	Sequence 1, Appl
C 58	8	2.4	45136	3	US-09-422-869-1	Sequence 1, Appl
C 59	8	2.4	54245	4	US-09-949-016-13499	Sequence 13499, A
C 60	8	2.4	134987	4	US-09-949-016-15348	Sequence 15348, A
C 61	8	2.4	134987	4	US-09-949-016-15349	Sequence 15349, A
C 62	8	2.4	134987	4	US-09-949-016-15350	Sequence 15350, A
C 63	8	2.4	134987	4	US-09-949-016-15507	Sequence 15507, A
C 64	8	2.4	134987	4	US-09-949-016-15508	Sequence 15508, A
C 65	8	2.4	134987	4	US-09-949-016-15509	Sequence 15509, A
C 66	8	2.4	146307	4	US-09-949-016-14881	Sequence 14881, A
C 67	8	2.4	146307	4	US-09-949-016-14882	Sequence 14882, A
C 68	8	2.4	146307	4	US-09-949-016-14883	Sequence 14883, A
C 69	8	2.4	146307	4	US-09-949-016-14884	Sequence 14884, A
C 70	8	2.4	146307	4	US-09-949-016-14885	Sequence 14885, A
C 71	8	2.4	146307	4	US-09-949-016-14886	Sequence 14886, A
C 72	8	2.4	146307	4	US-09-949-016-14887	Sequence 14887, A
C 73	8	2.4	146307	4	US-09-949-016-14888	Sequence 14888, A
C 74	8	2.4	184805	4	US-09-949-016-11747	Sequence 11747, A
C 75	8	2.4	184805	4	US-09-949-016-12835	Sequence 12835, A
C 76	8	2.4	184805	4	US-09-949-016-12836	Sequence 12836, A
C 77	8	2.4	184805	4	US-09-949-016-12837	Sequence 12837, A
C 78	8	2.4	175236	4	US-09-949-016-12833	Sequence 12833, A
C 79	8	2.4	177293	4	US-09-949-016-16513	Sequence 16513, A
C 80	8	2.4	187169	4	US-09-949-016-12776	Sequence 12776, A
C 81	8	2.4	191569	4	US-09-949-016-15940	Sequence 15940, A
C 82	8	2.4	256171	4	US-09-949-016-12822	Sequence 12822, A
C 83	8	2.4	256171	4	US-09-949-016-15524	Sequence 15524, A
C 84	8	2.4	4403765	3	US-09-103-840A-2	Sequence 2, Appl
C 85	8	2.4	4403765	3	US-09-103-840A-2	Sequence 2, Appl

86	8	2.4	4411529	3	US-09-103-840A-1	Sequence 1, Appli	7	2.1	498	4	US-09-252-991A-15197	Sequence 15197, A
C 87	8	2.4	4411529	3	US-09-103-840A-1	Sequence 1, Appli	7	2.1	503	4	US-09-621-976-17139	Sequence 17139, A
C 88	7	2.1	21	4	US-09-422-978-11680	Sequence 11680, A	7	2.1	504	4	US-09-248-796A-5217	Sequence 5217, Ap
C 89	7	2.1	25	4	US-09-396-196G-64569	Sequence 64569, A	7	2.1	506	4	US-09-513-999C-9805	Sequence 9805, Ap
C 90	7	2.1	97	5	PCT-US94-06456-26	Sequence 26, Appl	7	2.1	514	4	US-09-270-767-7141	Sequence 7141, Ap
C 91	7	2.1	97	5	PCT-US94-06456-55	Sequence 55, Appl	7	2.1	514	4	US-09-270-767-7141	Sequence 7141, Ap
C 92	7	2.1	147	4	US-09-328-352-3227	Sequence 3227, Ap	7	2.1	533	6	5482709-5	Patent No. 5482709
C 93	7	2.1	183	4	US-09-248-796A-12518	Sequence 12518, A	7	2.1	533	6	5482709-5	Patent No. 5482709
C 94	7	2.1	192	4	US-09-513-999C-16968	Sequence 16968, A	7	2.1	543	6	5273901-6	Patent No. 5273901
C 95	7	2.1	194	4	US-09-513-999C-12486	Sequence 12486, A	7	2.1	543	6	5273901-6	Patent No. 5273901
C 96	7	2.1	207	4	US-09-902-540-4076	Sequence 4076, Ap	7	2.1	553	3	US-09-513-999C-180	Sequence 180, App
C 97	7	2.1	207	4	US-09-248-796A-9966	Sequence 9966, Ap	7	2.1	553	3	US-09-385-982-431	Sequence 431, App
C 98	7	2.1	210	4	US-09-421-338A-96	Sequence 96, Appl	7	2.1	565	3	US-09-710-279-2657	Sequence 2697, Ap
C 99	7	2.1	220	4	US-09-021-338A-96	Sequence 96, Appl	7	2.1	567	4	US-09-248-796A-6399	Sequence 6399, Ap
C 100	7	2.1	231	4	US-09-107-433-440	Sequence 433, Appl	7	2.1	573	4	US-09-134-001C-2549	Sequence 2549, Ap
C 101	7	2.1	234	4	US-09-621-976-4117	Sequence 4117, Ap	7	2.1	579	3	US-09-248-796A-13766	Sequence 13766, A
C 102	7	2.1	237	4	US-09-489-039A-2474	Sequence 2474, Ap	7	2.1	582	4	US-09-252-991A-8856	Sequence 8856, Ap
C 103	7	2.1	239	4	US-09-489-039A-2474	Sequence 2474, Ap	7	2.1	584	4	US-09-543-681A-2177	Sequence 2177, Ap
C 104	7	2.1	240	4	US-09-513-999C-33598	Sequence 33598, A	7	2.1	594	4	US-09-543-681A-2177	Sequence 2177, Ap
C 105	7	2.1	243	3	US-09-513-999C-33598	Sequence 33598, A	7	2.1	597	4	US-09-598-401C-100	Sequence 100, Appl
C 106	7	2.1	272	4	US-09-313-294A-4117	Sequence 4117, Ap	7	2.1	600	4	US-09-408-020-21	Sequence 21, Appl
C 107	7	2.1	277	4	US-09-489-039A-4520	Sequence 4520, Ap	7	2.1	601	4	US-09-949-016-19666	Sequence 19666, A
C 108	7	2.1	282	4	US-09-489-039A-4520	Sequence 4520, Ap	7	2.1	601	4	US-09-949-016-21191	Sequence 21191, A
C 109	7	2.1	292	4	US-09-513-999C-27614	Sequence 27614, A	7	2.1	601	4	US-09-949-016-27815	Sequence 27815, A
C 110	7	2.1	294	4	US-09-543-681A-2569	Sequence 2569, Ap	7	2.1	601	4	US-09-949-016-32695	Sequence 32695, A
C 111	7	2.1	306	4	US-09-252-991A-14703	Sequence 14703, A	7	2.1	601	4	US-09-949-016-35342	Sequence 35342, A
C 112	7	2.1	306	4	US-09-328-352-2197	Sequence 2197, Ap	7	2.1	601	4	US-09-949-016-47058	Sequence 47058, A
C 113	7	2.1	312	4	US-09-489-039A-3085	Sequence 3085, Ap	7	2.1	601	4	US-09-949-016-52133	Sequence 52133, A
C 114	7	2.1	324	1	US-08-592-126-120	Sequence 126, App	7	2.1	601	4	US-09-949-016-53986	Sequence 53986, A
C 115	7	2.1	324	4	US-09-168-595-120	Sequence 120, App	7	2.1	601	4	US-09-949-016-55127	Sequence 55127, A
C 116	7	2.1	325	4	US-09-513-999C-35228	Sequence 35228, A	7	2.1	601	4	US-09-949-016-65234	Sequence 65234, A
C 117	7	2.1	328	4	US-09-471-276-254	Sequence 254, App	7	2.1	601	4	US-09-949-016-69685	Sequence 69685, A
C 118	7	2.1	332	4	US-09-513-999C-1578	Sequence 1578, Ap	7	2.1	601	4	US-09-949-016-70408	Sequence 70408, A
C 119	7	2.1	363	4	US-09-902-540-3936	Sequence 3936, Ap	7	2.1	601	4	US-09-949-016-73301	Sequence 73301, A
C 120	7	2.1	373	3	US-09-280-116-25	Sequence 25, Appl	7	2.1	601	4	US-09-949-016-74851	Sequence 74851, A
C 121	7	2.1	379	1	US-08-145-617-5	Sequence 5, Appl	7	2.1	601	4	US-09-949-016-75050	Sequence 75050, A
C 122	7	2.1	384	4	US-09-902-540-7527	Sequence 7527, Ap	7	2.1	601	4	US-09-949-016-75249	Sequence 75249, A
C 123	7	2.1	387	4	US-09-513-999C-2236	Sequence 2236, Ap	7	2.1	601	4	US-09-949-016-75448	Sequence 75448, A
C 124	7	2.1	399	4	US-09-583-110-1470	Sequence 1470, Ap	7	2.1	601	4	US-09-949-016-75647	Sequence 75647, A
C 125	7	2.1	400	4	US-09-355-709C-6	Sequence 6, Appl	7	2.1	601	4	US-09-949-016-76149	Sequence 76149, A
C 126	7	2.1	400	4	US-08-956-171E-3743	Sequence 3743, Ap	7	2.1	601	4	US-09-949-016-81554	Sequence 81554, A
C 127	7	2.1	400	4	US-08-956-171E-4051	Sequence 4051, Ap	7	2.1	601	4	US-09-949-016-81555	Sequence 81555, A
C 128	7	2.1	400	4	US-08-781-986A-3743	Sequence 3743, Ap	7	2.1	601	4	US-09-949-016-81557	Sequence 81557, A
C 129	7	2.1	400	4	US-08-781-986A-4051	Sequence 4051, Ap	7	2.1	601	4	US-09-949-016-84956	Sequence 84956, A
C 130	7	2.1	414	4	US-09-252-991A-8746	Sequence 8746, Ap	7	2.1	601	4	US-09-949-016-86791	Sequence 86791, A
C 131	7	2.1	414	4	US-09-107-433-1001	Sequence 1001, Ap	7	2.1	601	4	US-09-949-016-95315	Sequence 95315, A
C 132	7	2.1	422	4	US-09-621-976-8754	Sequence 8754, Ap	7	2.1	601	4	US-09-949-016-95493	Sequence 95493, A
C 133	7	2.1	423	1	US-08-470-179-66	Sequence 66, Appl	7	2.1	601	4	US-09-949-016-95671	Sequence 95671, A
C 134	7	2.1	423	1	US-08-470-179-67	Sequence 67, Appl	7	2.1	601	4	US-09-949-016-95849	Sequence 95849, A
C 135	7	2.1	432	3	US-09-397-787-207	Sequence 207, App	7	2.1	601	4	US-09-949-016-106205	Sequence 106205, A
C 136	7	2.1	432	3	US-09-252-991A-3546	Sequence 3546, Ap	7	2.1	601	4	US-09-949-016-106404	Sequence 106404, A
C 137	7	2.1	435	4	US-09-248-796A-8821	Sequence 8821, Ap	7	2.1	601	4	US-09-949-016-106603	Sequence 106603, A
C 138	7	2.1	440	4	US-09-270-767-3311	Sequence 3311, Ap	7	2.1	601	4	US-09-949-016-106603	Sequence 106603, A
C 139	7	2.1	440	4	US-09-270-767-18593	Sequence 18593, A	7	2.1	601	4	US-09-949-016-106802	Sequence 106802, A
C 140	7	2.1	440	4	US-09-513-999C-648	Sequence 648, App	7	2.1	601	4	US-09-949-016-107001	Sequence 107001, A
C 141	7	2.1	441	4	US-09-643-597-253	Sequence 253, App	7	2.1	601	4	US-09-949-016-107043	Sequence 107043, A
C 142	7	2.1	462	3	US-09-480-884A-253	Sequence 253, App	7	2.1	601	4	US-09-949-016-107085	Sequence 107085, A
C 143	7	2.1	462	4	US-09-542-615A-253	Sequence 253, App	7	2.1	601	4	US-09-949-016-107127	Sequence 107127, A
C 144	7	2.1	462	4	US-09-606-421B-253	Sequence 253, App	7	2.1	601	4	US-09-949-016-107169	Sequence 107169, A
C 145	7	2.1	462	4	US-09-476-496A-253	Sequence 253, App	7	2.1	601	4	US-09-949-016-107211	Sequence 107211, A
C 146	7	2.1	462	4	US-09-630-940B-253	Sequence 253, App	7	2.1	601	4	US-09-949-016-107253	Sequence 107253, A
C 147	7	2.1	464	4	US-09-621-976-14241	Sequence 14241, A	7	2.1	601	4	US-09-949-016-126784	Sequence 126784, A
C 148	7	2.1	465	4	US-09-252-991A-8912	Sequence 8912, Ap	7	2.1	601	4	US-09-949-016-126833	Sequence 126833, A
C 149	7	2.1	482	4	US-09-702-705-843	Sequence 843, App	7	2.1	601	4	US-09-949-016-134417	Sequence 134417, A
C 150	7	2.1	482	4	US-09-736-457-843	Sequence 843, App	7	2.1	601	4	US-09-949-016-134465	Sequence 134465, A
C 151	7	2.1	482	4	US-09-614-124B-843	Sequence 843, App	7	2.1	601	4	US-09-949-016-134515	Sequence 134515, A
C 152	7	2.1	482	4	US-09-671-325-843	Sequence 843, App	7	2.1	601	4	US-09-949-016-135704	Sequence 135704, A
C 153	7	2.1	482	4	US-09-658-824-843	Sequence 843, App	7	2.1	601	4	US-09-949-016-142415	Sequence 142415, A
C 154	7	2.1	486	4	US-09-270-767-2618	Sequence 2618, Ap	7	2.1	601	4	US-09-949-016-144728	Sequence 144728, A
C 155	7	2.1	486	4	US-09-270-767-17900	Sequence 17900, A	7	2.1	601	4	US-09-949-016-147690	Sequence 147690, A
C 156	7	2.1	486	4	US-09-639-207-13	Sequence 13, Appl	7	2.1	601	4	US-09-949-016-147691	Sequence 147691, A
C 157	7	2.1	489	4	US-09-248-796A-7305	Sequence 7305, Ap	7	2.1	601	4	US-09-949-016-150281	Sequence 150281, A
C 158	7	2.1	489	4	US-09-248-796A-7305	Sequence 7305, Ap	7	2.1	601	4	US-09-949-016-150281	Sequence 150281, A

C 232	7	2.1	601	4	US-09-949-016-150328	Sequence 150328,	305	7	2.1	935	4	US-09-513-999C-967	Sequence 967, App
C 233	7	2.1	601	4	US-09-949-016-153691	Sequence 153691,	306	7	2.1	942	4	US-09-502-540-2655	Sequence 2655, Ap
C 234	7	2.1	601	4	US-09-949-016-155308	Sequence 155308,	307	7	2.1	948	4	US-09-252-991A-9121	Sequence 9121, Ap
C 235	7	2.1	601	4	US-09-949-016-155334	Sequence 155334,	308	7	2.1	951	4	US-09-543-681A-1701	Sequence 1701, Ap
C 236	7	2.1	601	4	US-09-949-016-158632	Sequence 158632,	C 309	7	2.1	963	4	US-09-710-279-1785	Sequence 1785, Ap
C 237	7	2.1	601	4	US-09-949-016-160601	Sequence 160601,	310	7	2.1	966	4	US-09-252-991A-15581	Sequence 15581, A
C 238	7	2.1	601	4	US-09-949-016-160602	Sequence 160602,	311	7	2.1	975	3	US-09-071-035-37	Sequence 37, Appl
C 239	7	2.1	601	4	US-09-949-016-162287	Sequence 162287,	312	7	2.1	978	4	US-09-543-681A-1393	Sequence 1393, Ap
C 240	7	2.1	601	4	US-09-949-016-162288	Sequence 162288,	C 313	7	2.1	978	4	US-09-248-796A-5978	Sequence 5978, Ap
C 241	7	2.1	601	4	US-09-949-016-166487	Sequence 166487,	314	7	2.1	984	4	US-09-134-000C-937	Sequence 937, App
C 242	7	2.1	601	4	US-09-949-016-168100	Sequence 168100,	C 315	7	2.1	986	4	US-09-270-767-2302	Sequence 2302, Ap
C 243	7	2.1	601	4	US-09-949-016-169415	Sequence 169415,	C 316	7	2.1	986	4	US-09-270-767-17584	Sequence 17584, A
C 244	7	2.1	601	4	US-09-949-016-169416	Sequence 169416,	C 317	7	2.1	993	3	US-09-134-001C-1788	Sequence 1788, Ap
C 245	7	2.1	601	4	US-09-949-016-169417	Sequence 169417,	318	7	2.1	998	3	US-08-858-207A-150	Sequence 150, App
C 246	7	2.1	601	4	US-09-949-016-170440	Sequence 170440,	319	7	2.1	1008	4	US-09-248-796A-2996	Sequence 2996, Ap
C 247	7	2.1	601	4	US-09-949-016-170441	Sequence 170441,	C 320	7	2.1	1024	4	US-09-328-475C-85	Sequence 85, Appl
C 248	7	2.1	601	4	US-09-949-016-175593	Sequence 175593,	321	7	2.1	1038	4	US-09-248-796A-456	Sequence 456, App
C 249	7	2.1	601	4	US-09-949-016-177594	Sequence 177594,	322	7	2.1	1038	4	US-10-121-857-70	Sequence 70, Appl
C 250	7	2.1	601	4	US-09-949-016-177594	Sequence 177594,	323	7	2.1	1053	4	US-09-489-039A-2261	Sequence 2261, Ap
C 251	7	2.1	601	4	US-09-949-016-177857	Sequence 177857,	324	7	2.1	1072	4	US-09-270-767-11001	Sequence 11001, A
C 252	7	2.1	601	4	US-09-949-016-179109	Sequence 179109,	C 325	7	2.1	1089	4	US-09-270-767-10641	Sequence 10641, A
C 253	7	2.1	601	4	US-09-949-016-179110	Sequence 179110,	C 326	7	2.1	1095	4	US-09-248-796A-1700	Sequence 1700, Ap
C 254	7	2.1	601	4	US-09-949-016-180642	Sequence 180642,	C 327	7	2.1	1104	4	US-09-543-681A-4027	Sequence 4027, Ap
C 255	7	2.1	601	4	US-09-949-016-185772	Sequence 185772,	C 328	7	2.1	1122	4	US-09-252-991A-8798	Sequence 8798, Ap
C 256	7	2.1	601	4	US-09-949-016-186879	Sequence 186879,	C 329	7	2.1	1131	4	US-09-107-532A-2794	Sequence 2794, Ap
C 257	7	2.1	601	4	US-09-949-016-190164	Sequence 190164,	C 330	7	2.1	1135	4	US-09-203-959A-3	Sequence 3, Appli
C 258	7	2.1	601	4	US-09-949-016-190165	Sequence 190165,	C 331	7	2.1	1143	4	US-09-270-767-12606	Sequence 12606, A
C 259	7	2.1	601	4	US-09-949-016-191608	Sequence 191608,	332	7	2.1	1161	4	US-09-502-540-6414	Sequence 6414, Ap
C 260	7	2.1	601	4	US-09-949-016-191785	Sequence 191785,	C 333	7	2.1	1182	4	US-09-328-352-3041	Sequence 3041, Ap
C 261	7	2.1	601	4	US-09-949-016-191964	Sequence 191964,	C 334	7	2.1	1185	4	US-09-328-352-2675	Sequence 2675, Ap
C 262	7	2.1	601	4	US-09-949-016-192142	Sequence 192142,	C 335	7	2.1	1194	4	US-09-489-039A-160	Sequence 160, App
C 263	7	2.1	601	4	US-09-949-016-200196	Sequence 200196,	C 336	7	2.1	1202	4	US-09-949-016-2208	Sequence 2208, Ap
C 264	7	2.1	602	4	US-09-636-215-767	Sequence 200197,	C 337	7	2.1	1203	4	US-09-489-039A-4745	Sequence 4745, Ap
C 265	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 338	7	2.1	1215	4	US-09-220-132-184	Sequence 184, App
C 266	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 339	7	2.1	1218	4	US-09-252-991A-9375	Sequence 9375, Ap
C 267	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 340	7	2.1	1224	4	US-09-270-767-292	Sequence 292, App
C 268	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 341	7	2.1	1224	4	US-09-270-767-15574	Sequence 15574, A
C 269	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 342	7	2.1	1234	4	US-09-620-312D-232	Sequence 232, App
C 270	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 343	7	2.1	1234	4	US-09-495-050A-187	Sequence 187, App
C 271	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 344	7	2.1	1253	4	US-09-949-016-1511	Sequence 1511, Ap
C 272	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 345	7	2.1	1253	4	US-09-949-016-1512	Sequence 1512, Ap
C 273	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 346	7	2.1	1266	4	US-09-489-039A-812	Sequence 812, App
C 274	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 347	7	2.1	1279	4	US-09-270-767-14878	Sequence 14878, A
C 275	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 348	7	2.1	1284	3	US-08-985-526-24	Sequence 24, Appl
C 276	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 349	7	2.1	1287	4	US-09-248-796A-4715	Sequence 4715, Ap
C 277	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 350	7	2.1	1341	4	US-10-162-012-7	Sequence 7, Appli
C 278	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 351	7	2.1	1348	3	US-09-152-060-14	Sequence 14, Appli
C 279	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 352	7	2.1	1363	4	US-09-653-375B-1	Sequence 1, Appli
C 280	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 353	7	2.1	1373	4	US-09-799-451-708	Sequence 708, App
C 281	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 354	7	2.1	1388	4	US-09-655-908-1	Sequence 1, Appli
C 282	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 355	7	2.1	1391	4	US-09-328-352-1856	Sequence 13275, A
C 283	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 356	7	2.1	1392	4	US-09-328-352-1856	Sequence 1856, Ap
C 284	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 357	7	2.1	1394	4	US-09-774-528-4	Sequence 4, Appli
C 285	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 358	7	2.1	1398	4	US-09-248-796A-961	Sequence 961, App
C 286	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 359	7	2.1	1413	3	US-09-813-918-1	Sequence 1, Appli
C 287	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 360	7	2.1	1413	4	US-10-060-311-1	Sequence 1, Appli
C 288	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 361	7	2.1	1425	3	US-08-936-165A-254	Sequence 254, App
C 289	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 362	7	2.1	1434	4	US-09-252-991A-8967	Sequence 8967, Ap
C 290	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 363	7	2.1	1440	4	US-09-248-796A-1804	Sequence 1804, Ap
C 291	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 364	7	2.1	1440	4	US-09-252-991A-3554	Sequence 3554, Ap
C 292	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 365	7	2.1	1443	4	US-09-248-796A-4907	Sequence 4907, Ap
C 293	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 366	7	2.1	1443	4	US-09-489-039A-3189	Sequence 3189, Ap
C 294	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 367	7	2.1	1452	4	US-09-370-838-80	Sequence 80, Appl
C 295	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 368	7	2.1	1460	4	US-09-854-133-80	Sequence 80, Appl
C 296	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 369	7	2.1	1461	4	US-09-540-236-1593	Sequence 1593, Ap
C 297	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 370	7	2.1	1462	3	US-09-370-838-55	Sequence 55, Appl
C 298	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 371	7	2.1	1462	4	US-09-854-133-55	Sequence 55, Appl
C 299	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 372	7	2.1	1488	4	US-09-252-991A-4168	Sequence 4168, Ap
C 300	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 373	7	2.1	1503	4	US-09-252-991A-4643	Sequence 4643, Ap
C 301	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 374	7	2.1	1517	1	US-08-508-448C-15	Sequence 15, Appl
C 302	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 375	7	2.1	1524	4	US-09-949-016-5791	Sequence 5791, Ap
C 303	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 376	7	2.1	1545	3	US-09-292-858B-13	Sequence 13, Appl
C 304	7	2.1	602	4	US-09-685-168A-767	Sequence 767, App	C 377	7	2.1	1545	3		

C 378	7	2.1	1563	3	US-09-292-858B-11	Sequence 11, Appl	451	7	2.1	1959	3	US-08-996-441B-43	Sequence 43, Appl
C 379	7	2.1	1584	4	US-08-107-532A-1168	Sequence 1168, Ap	452	7	2.1	1959	3	US-08-996-441B-45	Sequence 45, Appl
C 380	7	2.1	1584	4	US-09-107-532A-1924	Sequence 1924, Ap	453	7	2.1	1959	3	US-08-996-441B-47	Sequence 47, Appl
C 381	7	2.1	1585	4	US-09-023-655-1329	Sequence 1329, Ap	454	7	2.1	1959	3	US-08-996-441B-49	Sequence 49, Appl
C 382	7	2.1	1632	4	US-09-252-991A-15621	Sequence 15621, A	455	7	2.1	1959	3	US-08-996-441B-53	Sequence 53, Appl
C 383	7	2.1	1641	4	US-09-489-039A-2456	Sequence 2456, Ap	456	7	2.1	1959	3	US-08-996-441B-59	Sequence 59, Appl
C 384	7	2.1	1655	4	US-09-248-796A-6368	Sequence 6368, Ap	457	7	2.1	1959	3	US-08-996-441B-61	Sequence 61, Appl
C 385	7	2.1	1683	4	US-09-552-991A-9428	Sequence 9428, Ap	458	7	2.1	1959	3	US-08-996-441B-63	Sequence 63, Appl
C 386	7	2.1	1689	4	US-09-902-540-2528	Sequence 2528, Ap	459	7	2.1	1959	3	US-08-996-441B-65	Sequence 65, Appl
C 387	7	2.1	1719	3	US-09-160-065-1	Sequence 1, Appl1	460	7	2.1	1959	3	US-08-996-441B-67	Sequence 67, Appl
C 388	7	2.1	1722	4	US-09-902-540-3804	Sequence 3804, Ap	461	7	2.1	1959	3	US-08-996-441B-97	Sequence 97, Appl
C 389	7	2.1	1770	4	US-08-976-063B-29	Sequence 29, Appl	462	7	2.1	1959	3	US-08-993-722A-1	Sequence 1, Appl1
C 390	7	2.1	1791	4	US-09-252-991A-9460	Sequence 9460, Ap	463	7	2.1	1959	3	US-08-993-722A-3	Sequence 3, Appl1
C 391	7	2.1	1824	4	US-09-540-236-86	Sequence 86, Appl	464	7	2.1	1959	3	US-08-993-722A-5	Sequence 5, Appl1
C 392	7	2.1	1830	4	US-09-252-991A-3563	Sequence 3563, Ap	465	7	2.1	1959	3	US-08-993-722A-7	Sequence 7, Appl1
C 393	7	2.1	1838	3	US-09-227-357-32	Sequence 3, Appl	466	7	2.1	1959	3	US-08-993-722A-9	Sequence 9, Appl1
C 394	7	2.1	1843	3	US-08-718-738-1	Sequence 1, Appl1	467	7	2.1	1959	3	US-08-993-722A-11	Sequence 11, Appl
C 395	7	2.1	1843	3	US-09-221-844-1	Sequence 1, Appl1	468	7	2.1	1959	3	US-08-993-722A-13	Sequence 13, Appl
C 396	7	2.1	1843	4	US-09-843-846-1	Sequence 1, Appl1	469	7	2.1	1959	3	US-08-993-722A-15	Sequence 15, Appl
C 397	7	2.1	1843	5	FCT-US95-0332A-1	Sequence 11207, A	470	7	2.1	1959	3	US-08-993-722A-17	Sequence 17, Appl
C 398	7	2.1	1845	4	US-09-252-991A-11207	Sequence 467, App	471	7	2.1	1959	3	US-08-993-722A-19	Sequence 19, Appl
C 399	7	2.1	1864	3	US-09-071-035-467	Sequence 11062, A	472	7	2.1	1959	3	US-08-993-722A-21	Sequence 21, Appl
C 400	7	2.1	1891	4	US-09-270-767-11062	Sequence 1429, Ap	473	7	2.1	1959	3	US-08-993-722A-23	Sequence 23, Appl
C 401	7	2.1	1896	4	US-09-252-991A-1429	Sequence 50, Appl	474	7	2.1	1959	3	US-08-993-722A-25	Sequence 25, Appl
C 402	7	2.1	1908	3	US-08-845-258-50	Sequence 50, Appl	475	7	2.1	1959	3	US-08-993-722A-27	Sequence 27, Appl
C 403	7	2.1	1908	3	US-08-950-571-50	Sequence 50, Appl	476	7	2.1	1959	3	US-08-993-722A-29	Sequence 29, Appl
C 404	7	2.1	1908	3	US-09-528-784A-50	Sequence 50, Appl	477	7	2.1	1959	3	US-08-993-722A-31	Sequence 31, Appl
C 405	7	2.1	1908	4	US-09-569-098A-50	Sequence 1, Appl1	478	7	2.1	1959	3	US-08-993-722A-33	Sequence 33, Appl
C 406	7	2.1	1920	1	US-08-186-222-1	Sequence 1, Appl1	479	7	2.1	1959	3	US-08-993-722A-35	Sequence 35, Appl
C 407	7	2.1	1926	4	US-09-252-991A-3531	Sequence 5558, Ap	480	7	2.1	1959	3	US-08-993-722A-37	Sequence 37, Appl
C 408	7	2.1	1938	4	US-08-949-016-5558	Sequence 5, Appl1	481	7	2.1	1959	3	US-08-993-722A-39	Sequence 39, Appl
C 409	7	2.1	1953	1	US-08-315-468-5	Sequence 3, Appl1	482	7	2.1	1959	3	US-08-993-722A-41	Sequence 41, Appl
C 410	7	2.1	1953	3	US-07-941-650A-3	Sequence 770, App	483	7	2.1	1959	3	US-08-993-722A-43	Sequence 43, Appl
C 411	7	2.1	1955	4	US-09-949-016-770	Sequence 51, Appl	484	7	2.1	1959	3	US-08-993-722A-45	Sequence 45, Appl
C 412	7	2.1	1956	3	US-08-996-441B-51	Sequence 55, Appl	485	7	2.1	1959	3	US-08-993-722A-47	Sequence 47, Appl
C 413	7	2.1	1956	3	US-08-996-441B-55	Sequence 55, Appl	486	7	2.1	1959	3	US-08-993-722A-49	Sequence 49, Appl
C 414	7	2.1	1956	3	US-08-996-441B-57	Sequence 51, Appl	487	7	2.1	1959	3	US-08-993-722A-51	Sequence 51, Appl
C 415	7	2.1	1956	3	US-08-993-722A-51	Sequence 55, Appl	488	7	2.1	1959	3	US-08-993-722A-53	Sequence 53, Appl
C 416	7	2.1	1956	3	US-08-993-722A-55	Sequence 55, Appl	489	7	2.1	1959	3	US-08-993-722A-55	Sequence 55, Appl
C 417	7	2.1	1956	3	US-08-993-722A-57	Sequence 51, Appl	490	7	2.1	1959	3	US-08-993-722A-59	Sequence 59, Appl
C 418	7	2.1	1956	3	US-08-993-170A-51	Sequence 51, Appl	491	7	2.1	1959	3	US-08-993-722A-61	Sequence 61, Appl
C 419	7	2.1	1956	3	US-08-993-170A-55	Sequence 55, Appl	492	7	2.1	1959	3	US-08-993-722A-63	Sequence 63, Appl
C 420	7	2.1	1956	3	US-08-993-170A-57	Sequence 55, Appl	493	7	2.1	1959	3	US-08-993-722A-65	Sequence 65, Appl
C 421	7	2.1	1956	3	US-08-993-775B-51	Sequence 51, Appl	494	7	2.1	1959	3	US-08-993-722A-67	Sequence 67, Appl
C 422	7	2.1	1956	3	US-08-993-775B-55	Sequence 55, Appl	495	7	2.1	1959	3	US-08-993-722A-69	Sequence 69, Appl
C 423	7	2.1	1956	3	US-08-993-775B-57	Sequence 57, Appl	496	7	2.1	1959	3	US-08-993-722A-71	Sequence 71, Appl
C 424	7	2.1	1956	4	US-09-427-770-51	Sequence 51, Appl	497	7	2.1	1959	3	US-08-993-722A-73	Sequence 73, Appl
C 425	7	2.1	1956	4	US-09-427-770-55	Sequence 55, Appl	498	7	2.1	1959	3	US-08-993-170A-5	Sequence 5, Appl1
C 426	7	2.1	1956	4	US-09-427-770-57	Sequence 57, Appl	499	7	2.1	1959	3	US-08-993-170A-7	Sequence 7, Appl1
C 427	7	2.1	1956	4	US-09-427-769-51	Sequence 51, Appl	500	7	2.1	1959	3	US-08-993-170A-9	Sequence 9, Appl1
C 428	7	2.1	1956	4	US-09-427-769-55	Sequence 55, Appl	501	7	2.1	1959	3	US-08-993-170A-11	Sequence 11, Appl
C 429	7	2.1	1956	4	US-08-996-441B-1	Sequence 1, Appl1	502	7	2.1	1959	3	US-08-993-170A-13	Sequence 13, Appl
C 430	7	2.1	1959	3	US-08-996-441B-5	Sequence 5, Appl1	503	7	2.1	1959	3	US-08-993-170A-15	Sequence 15, Appl
C 431	7	2.1	1959	3	US-08-996-441B-5	Sequence 7, Appl1	504	7	2.1	1959	3	US-08-993-170A-17	Sequence 17, Appl
C 432	7	2.1	1959	3	US-08-996-441B-7	Sequence 9, Appl1	505	7	2.1	1959	3	US-08-993-170A-19	Sequence 19, Appl
C 433	7	2.1	1959	3	US-08-996-441B-9	Sequence 11, Appl	506	7	2.1	1959	3	US-08-993-170A-21	Sequence 21, Appl
C 434	7	2.1	1959	3	US-08-996-441B-11	Sequence 13, Appl	507	7	2.1	1959	3	US-08-993-170A-23	Sequence 23, Appl
C 435	7	2.1	1959	3	US-08-996-441B-13	Sequence 15, Appl	508	7	2.1	1959	3	US-08-993-170A-25	Sequence 25, Appl
C 436	7	2.1	1959	3	US-08-996-441B-15	Sequence 17, Appl	509	7	2.1	1959	3	US-08-993-170A-27	Sequence 27, Appl
C 437	7	2.1	1959	3	US-08-996-441B-17	Sequence 19, Appl	510	7	2.1	1959	3	US-08-993-170A-29	Sequence 29, Appl
C 438	7	2.1	1959	3	US-08-996-441B-19	Sequence 21, Appl	511	7	2.1	1959	3	US-08-993-170A-31	Sequence 31, Appl
C 439	7	2.1	1959	3	US-08-996-441B-21	Sequence 23, Appl	512	7	2.1	1959	3	US-08-993-170A-33	Sequence 33, Appl
C 440	7	2.1	1959	3	US-08-996-441B-23	Sequence 25, Appl	513	7	2.1	1959	3	US-08-993-170A-35	Sequence 35, Appl
C 441	7	2.1	1959	3	US-08-996-441B-25	Sequence 27, Appl	514	7	2.1	1959	3	US-08-993-170A-37	Sequence 37, Appl
C 442	7	2.1	1959	3	US-08-996-441B-29	Sequence 29, Appl	515	7	2.1	1959	3	US-08-993-170A-39	Sequence 39, Appl
C 443	7	2.1	1959	3	US-08-996-441B-33	Sequence 31, Appl	516	7	2.1	1959	3	US-08-993-170A-41	Sequence 41, Appl
C 444	7	2.1	1959	3	US-08-996-441B-35	Sequence 33, Appl	517	7	2.1	1959	3	US-08-993-170A-43	Sequence 43, Appl
C 445	7	2.1	1959	3	US-08-996-441B-37	Sequence 35, Appl	518	7	2.1	1959	3	US-08-993-170A-45	Sequence 45, Appl
C 446	7	2.1	1959	3	US-08-996-441B-39	Sequence 37, Appl	519	7	2.1	1959	3	US-08-993-170A-47	Sequence 47, Appl
C 447	7	2.1	1959	3	US-08-996-441B-41	Sequence 39, Appl	520	7	2.1	1959	3	US-08-993-170A-49	Sequence 49, Appl
C 448	7	2.1	1959	3	US-08-996-441B-43	Sequence 41, Appl	521	7	2.1	1959	3	US-08-993-170A-51	Sequence 51, Appl
C 449	7	2.1	1959	3	US-08-996-441B-45	Sequence 43, Appl	522	7	2.1	1959	3	US-08-993-170A-53	Sequence 53, Appl
C 450	7	2.1	1959	3	US-08-996-441B-47	Sequence 45, Appl	523	7	2.1	1959	3	US-08-993-170A-55	Sequence 55, Appl

c 670	7	2.1	2091	4	US-09-620-312D-743	Sequence 743, App	c 743	7	2.1	3423	4	US-09-614-221A-91	Sequence 91, Appl
c 671	7	2.1	2097	3	US-09-134-001C-795	Sequence 795, App	c 744	7	2.1	3438	4	US-10-164-595-29	Sequence 29, Appl
c 672	7	2.1	2121	3	US-09-134-001C-2071	Sequence 2071, Ap	c 745	7	2.1	3450	4	US-09-377-466B-17	Sequence 17, Appl
c 673	7	2.1	2122	4	US-09-710-279-4457	Sequence 4457, Ap	c 746	7	2.1	3455	4	US-09-377-466B-36	Sequence 36, Appl
c 674	7	2.1	2126	4	US-09-620-312D-598	Sequence 598, App	c 747	7	2.1	3459	4	US-09-710-279-3553	Sequence 3553, Ap
c 675	7	2.1	2148	3	US-09-643-597-154	Sequence 154, App	c 748	7	2.1	3469	4	US-09-377-466B-23	Sequence 23, Appl
c 676	7	2.1	2148	4	US-09-480-884A-154	Sequence 154, App	c 749	7	2.1	3513	4	US-09-815-923-13	Sequence 13, Appl
c 677	7	2.1	2148	4	US-09-542-615A-154	Sequence 154, App	c 750	7	2.1	3522	4	US-09-543-681A-3793	Sequence 3793, Ap
c 678	7	2.1	2148	4	US-09-606-421B-154	Sequence 154, App	c 751	7	2.1	3532	2	US-09-205-204-1	Sequence 1, Appl
c 679	7	2.1	2148	4	US-09-221-107-154	Sequence 154, App	c 752	7	2.1	3549	4	US-09-710-279-3796	Sequence 3796, Ap
c 680	7	2.1	2148	4	US-09-466-396A-154	Sequence 154, App	c 753	7	2.1	3618	4	US-09-710-279-3443	Sequence 3443, Ap
c 681	7	2.1	2148	4	US-09-476-496A-154	Sequence 154, App	c 754	7	2.1	3627	4	US-09-710-279-3764	Sequence 3764, Ap
c 682	7	2.1	2148	4	US-09-630-940B-154	Sequence 154, App	c 755	7	2.1	3704	4	US-10-160-719A-57	Sequence 57, Appl
c 683	7	2.1	2148	4	US-09-285-479-154	Sequence 154, App	c 756	7	2.1	3732	3	US-09-212-971-7	Sequence 7, Appl
c 684	7	2.1	2156	4	US-09-086-663A-78	Sequence 78, Appl	c 757	7	2.1	3732	3	US-08-800-929A-7	Sequence 7, Appl
c 685	7	2.1	2157	3	US-08-973-005A-10	Sequence 10, Appl	c 758	7	2.1	3732	3	US-09-617-053A-7	Sequence 7, Appl
c 686	7	2.1	2158	4	US-09-949-016-3018	Sequence 3018, Ap	c 759	7	2.1	3754	4	US-09-377-466B-15	Sequence 15, Appl
c 687	7	2.1	2158	4	US-09-949-016-3026	Sequence 3026, Ap	c 760	7	2.1	3883	1	US-08-468-036-33	Sequence 33, Appl
c 688	7	2.1	2172	3	US-09-134-001C-2223	Sequence 2223, Ap	c 761	7	2.1	3883	2	US-08-376-843-33	Sequence 33, Appl
c 689	7	2.1	2202	4	US-09-902-540-8501	Sequence 8501, Ap	c 762	7	2.1	3884	3	US-09-541-782-3	Sequence 3, Appl
c 690	7	2.1	2253	4	US-09-949-016-3908	Sequence 3908, Ap	c 763	7	2.1	3884	3	US-10-270-085-3	Sequence 6, Appl
c 691	7	2.1	2256	4	US-09-949-016-2427	Sequence 2427, Ap	c 764	7	2.1	3901	2	US-08-574-959A-6	Sequence 4309, Ap
c 692	7	2.1	2256	4	US-09-949-016-2428	Sequence 2428, Ap	c 765	7	2.1	3901	3	US-09-357-014-6	Sequence 816, App
c 693	7	2.1	2280	3	US-08-996-441B-102	Sequence 102, App	c 766	7	2.1	3983	4	US-08-956-171E-816	Sequence 816, App
c 694	7	2.1	2280	3	US-08-993-722A-102	Sequence 102, App	c 767	7	2.1	3989	4	US-08-781-986A-816	Sequence 969, App
c 695	7	2.1	2280	3	US-08-993-775B-102	Sequence 102, App	c 768	7	2.1	3989	4	US-09-023-655-969	Sequence 229, App
c 696	7	2.1	2280	4	US-09-427-770-102	Sequence 102, App	c 769	7	2.1	4121	4	US-09-614-221A-229	Sequence 13, Appl
c 697	7	2.1	2280	4	US-09-427-769-102	Sequence 102, App	c 770	7	2.1	4146	4	US-09-377-466B-13	Sequence 236, App
c 698	7	2.1	2280	4	US-09-540-236-862	Sequence 862, App	c 771	7	2.1	4196	4	US-09-774-528-236	Sequence 400, App
c 699	7	2.1	2334	4	US-09-248-796A-1911	Sequence 1911, Ap	c 772	7	2.1	4209	4	US-09-248-796A-400	Sequence 4, Appl
c 700	7	2.1	2334	2	US-08-981-690-1	Sequence 1, Appl	c 773	7	2.1	4215	4	US-09-669-980A-6	Sequence 4, Appl
c 701	7	2.1	2364	4	US-09-252-991A-9040	Sequence 9040, Ap	c 774	7	2.1	4321	4	US-09-773-426A-4	Sequence 4, Appl
c 702	7	2.1	2370	4	US-09-252-991A-9252	Sequence 9252, Ap	c 775	7	2.1	4321	4	US-10-314-881-4	Sequence 546, App
c 703	7	2.1	2379	4	US-09-543-681A-2445	Sequence 2445, Ap	c 776	7	2.1	4336	4	US-08-956-171E-546	Sequence 546, App
c 704	7	2.1	2399	4	PCT-US92-00040-1	Sequence 2, Appl	c 777	7	2.1	4336	4	US-08-781-986A-546	Sequence 6, Appl
c 705	7	2.1	2430	5	US-09-618-425-4	Sequence 4, Appl	c 778	7	2.1	4336	4	US-08-809-254A-6	Sequence 15228, A
c 706	7	2.1	2430	4	US-09-774-528-223	Sequence 223, App	c 779	7	2.1	4528	4	US-09-949-016-15228	Sequence 9, Appl
c 707	7	2.1	2494	4	US-08-239-276-7	Sequence 7, Appl	c 780	7	2.1	4587	4	US-09-949-016-3235	Sequence 3235, Ap
c 708	7	2.1	2599	2	US-08-468-579B-7	Sequence 7, Appl	c 781	7	2.1	4621	4	US-09-231-899-76	Sequence 76, Appl
c 709	7	2.1	2599	3	US-08-468-577B-7	Sequence 7, Appl	c 782	7	2.1	4767	4	US-09-561-497-10	Sequence 10, Appl
c 710	7	2.1	2599	3	US-08-468-581A-7	Sequence 7, Appl	c 783	7	2.1	4793	3	US-09-270-767-10504	Sequence 10504, A
c 711	7	2.1	2599	4	US-09-205-258-110	Sequence 110, App	c 784	7	2.1	4860	4	US-09-445-353E-1	Sequence 5, Appl
c 712	7	2.1	2632	4	US-09-620-312D-474	Sequence 474, App	c 785	7	2.1	4899	2	US-08-210-535-5	Patent No. 5268270
c 713	7	2.1	2694	4	US-09-902-540-4257	Sequence 4257, Ap	c 786	7	2.1	4899	6	5268270-1	Patent No. 5268270
c 714	7	2.1	2735	4	US-09-370-838-79	Sequence 79, Appl	c 787	7	2.1	5164	4	US-09-949-016-16010	Sequence 16010, A
c 715	7	2.1	2790	3	US-09-854-133-79	Sequence 85, Appl	c 788	7	2.1	5203	4	US-09-799-451-613	Sequence 613, App
c 716	7	2.1	2900	1	US-08-592-136-85	Sequence 85, Appl	c 789	7	2.1	5203	4	US-09-949-016-13950	Sequence 13950, A
c 717	7	2.1	2995	1	US-09-168-595-85	Sequence 85, Appl	c 790	7	2.1	5215	4	US-09-231-899-71	Sequence 71, Appl
c 718	7	2.1	2995	4	US-09-710-279-3822	Sequence 3822, Ap	c 791	7	2.1	5215	4	US-09-943-016-4074	Sequence 4074, Ap
c 719	7	2.1	3036	4	US-09-377-466B-19	Sequence 19, Appl	c 792	7	2.1	5226	4	US-09-583-110-1314	Sequence 1314, Ap
c 720	7	2.1	3039	4	US-09-377-466B-21	Sequence 21, Appl	c 793	7	2.1	5304	4	US-09-107-433-1073	Sequence 1073, Ap
c 721	7	2.1	3039	4	US-09-377-466B-38	Sequence 38, Appl	c 794	7	2.1	5304	4	US-09-252-991A-1588	Sequence 1588, Ap
c 722	7	2.1	3044	4	US-09-620-312D-474	Sequence 474, App	c 795	7	2.1	5307	1	US-08-441-139-15	Sequence 1259, Ap
c 723	7	2.1	3063	4	US-09-270-767-13401	Sequence 13401, A	c 796	7	2.1	5338	4	US-09-949-016-1259	Sequence 1014, Ap
c 724	7	2.1	3074	4	US-08-551-437-3	Sequence 3, Appl	c 797	7	2.1	5522	4	US-09-442-100-1	Sequence 1, Appl
c 725	7	2.1	3084	1	US-09-004-225-3	Sequence 3, Appl	c 798	7	2.1	5523	4	US-08-939-106-1	Sequence 1, Appl
c 726	7	2.1	3084	3	US-09-084-346-3	Sequence 3, Appl	c 799	7	2.1	5720	3	US-09-442-100-1	Sequence 1, Appl
c 727	7	2.1	3084	3	US-09-104-704-3	Sequence 3, Appl	c 799	7	2.1	5720	4	US-09-976-594-183	Sequence 183, App
c 728	7	2.1	3084	4	US-09-902-540-536	Sequence 536, App	c 799	7	2.1	5933	4	US-09-949-016-17300	Sequence 17300, A
c 729	7	2.1	3084	4	US-08-943-768-1	Sequence 1, Appl	c 800	7	2.1	5939	4	US-09-949-016-12512	Sequence 12512, A
c 730	7	2.1	3150	3	US-09-865-960-1	Sequence 1, Appl	c 801	7	2.1	5941	4	US-09-949-016-869	Sequence 869, App
c 731	7	2.1	3150	4	US-08-050-684-1	Sequence 1, Appl	c 802	7	2.1	6019	4	US-09-949-016-2671	Sequence 2671, Ap
c 732	7	2.1	3175	1	US-08-582-719-1	Sequence 1, Appl	c 803	7	2.1	6019	4	US-09-134-000C-3207	Sequence 3207, Ap
c 733	7	2.1	3175	1	US-08-582-719-1	Sequence 1, Appl	c 804	7	2.1	6165	3	US-09-071-035-457	Sequence 457, App
c 734	7	2.1	3181	4	US-09-710-279-4118	Sequence 4118, Ap	c 805	7	2.1	6168	3	US-09-071-035-461	Sequence 461, App
c 735	7	2.1	3211	2	US-08-574-959A-8	Sequence 8, Appl	c 806	7	2.1	6168	3	US-09-071-035-465	Sequence 465, App
c 736	7	2.1	3211	3	US-09-357-014-8	Sequence 3492, Ap	c 807	7	2.1	6168	3		
c 737	7	2.1	3236	4	US-09-710-279-3492	Sequence 3492, Ap	c 808	7	2.1				
c 738	7	2.1	3297	4	US-09-620-312D-417	Sequence 417, App	c 809	7	2.1				
c 739	7	2.1	3376	1	US-08-320-559-29	Sequence 29, Appl	c 810	7	2.1				
c 740	7	2.1	3376	3	US-08-545-860D-29	Sequence 29, Appl	c 811	7	2.1				
c 741	7	2.1	3376	5	PCT-US94-04496-29	Sequence 29, Appl	c 812	7	2.1				
c 742	7	2.1	3376				c 813	7	2.1				
							c 814	7	2.1				
							c 815	7	2.1				

Sequence	Length	Matches	Conservative	Mismatches	Indels	Gaps
Sequence 14706, A	183	51	0	0	0	0
Sequence 14707, A	51	0	0	0	0	0
Sequence 14708, A	51	0	0	0	0	0
Sequence 14709, A	51	0	0	0	0	0
Sequence 15854, A	51	0	0	0	0	0
Sequence 12007, A	51	0	0	0	0	0
Sequence 16781, A	51	0	0	0	0	0
Sequence 12838, A	51	0	0	0	0	0
Sequence 17458, A	51	0	0	0	0	0
Sequence 11796, A	51	0	0	0	0	0
Sequence 12843, A	51	0	0	0	0	0
Sequence 12844, A	51	0	0	0	0	0
Sequence 12846, A	51	0	0	0	0	0
Sequence 13542, A	51	0	0	0	0	0
Sequence 13543, A	51	0	0	0	0	0
Sequence 13544, A	51	0	0	0	0	0
Sequence 13545, A	51	0	0	0	0	0
Sequence 14633, A	51	0	0	0	0	0
Sequence 14634, A	51	0	0	0	0	0
Sequence 14635, A	51	0	0	0	0	0
Sequence 14636, A	51	0	0	0	0	0
Sequence 14637, A	51	0	0	0	0	0
Sequence 14638, A	51	0	0	0	0	0
Sequence 14639, A	51	0	0	0	0	0
Sequence 14640, A	51	0	0	0	0	0
Sequence 12332, A	51	0	0	0	0	0
Sequence 30, Appl	51	0	0	0	0	0
Sequence 12150, A	51	0	0	0	0	0
Sequence 15374, A	51	0	0	0	0	0
Sequence 16910, A	51	0	0	0	0	0
Sequence 12175, A	51	0	0	0	0	0
Sequence 13638, A	51	0	0	0	0	0
Sequence 16057, A	51	0	0	0	0	0
Sequence 16076, A	51	0	0	0	0	0
Sequence 14712, A	51	0	0	0	0	0
Sequence 11803, A	51	0	0	0	0	0
Sequence 16739, A	51	0	0	0	0	0
Sequence 14817, A	51	0	0	0	0	0

; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 2556:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 242 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-956-171E-2556

Alignment Scores:
Pred. No.: 1,01e-17 Length: 242
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.76% Indels: 0
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-08-956-171E-2556 (1-242)

QY 102 TyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluSerLysLeuLysPro 121
Db 98 TATACCTCTGTAGGTACACGTAACAGCCAACTTAGAAGAAATTAGTAAATTAACCG 157
QY 122 AspleuilelleAlaAsp 127
Db 158 GATTTAATTATCGCTGAT 175

RESULT 4
US-08-781-986A-2556
; Sequence 2556, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:

US-10-724-972A-6352 (1-335) x US-08-956-171E-2556 (1-242)

QY 102 TyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluSerLysLeuLysPro 121
Db 98 TATACCTCTGTAGGTACACGTAACAGCCAACTTAGAAGAAATTAGTAAATTAACCG 157
QY 122 AspleuilelleAlaAsp 127
Db 158 GATTTAATTATCGCTGAT 175

RESULT 5
US-08-956-171E-355
; Sequence 355, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 2556:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 242 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-781-986A-2556

Alignment Scores:
Pred. No.: 1,01e-17 Length: 242
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.76% Indels: 0
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-08-781-986A-2556 (1-242)

QY 102 TyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluSerLysLeuLysPro 121
Db 98 TATACCTCTGTAGGTACACGTAACAGCCAACTTAGAAGAAATTAGTAAATTAACCG 157
QY 122 AspleuilelleAlaAsp 127
Db 158 GATTTAATTATCGCTGAT 175

RESULT 5

US-08-956-171E-355
; Sequence 355, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman

/ REGISTRATION NUMBER: 46,789
/ REFERENCE/DOCKET NUMBER: PB248P1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (240) 314-1224
/ TELEFAX: (301) 309-8439
/ INFORMATION FOR SEQ ID NO: 355:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 668 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 355:
US-08-956-171E-355

Alignment Scores: 3.21e-14 Length: 668
Pred. No.: 23.00 Matches: 23
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 6.87% Gaps: 0
DB: 4

US-10-724-972A-6352 (1-335) x US-08-781-986A-355 (1-668)
QY 306 TtpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeu 325
Db 290 TGGGCAAGATCTCGTGGCTTAATTTCTTGAAGAAATGCTTAAGAACTTTGTTGAATTA 349

QY 326 SerLysLys 328
Db 350 TCAAAAAA 358

RESULT 7
US-09-830-217-15
; Sequence 15, Application US/09830217
; Patent No. 6521441
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB461PCT
; CURRENT APPLICATION NUMBER: US/09/830,217
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: PCT/US99/06199
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/080,296
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/084,674
; PRIOR FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 999
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-830-217-15

Alignment Scores: 8.28e-05 Length: 999
Pred. No.: 14.00 Matches: 14
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 4.18% Gaps: 0
DB: 4

US-10-724-972A-6352 (1-335) x US-09-830-217-15 (1-999)
QY 111 ProAsnLeuGluGluIleSerLysLeuLysProAspLeuLeu 124
Db 322 CCTAACTTAGAGGAAATCTCTAAATTAACCGGACTTAAT 363

RESULT 8
US-10-278-946-15
; Sequence 15, Application US/10278946
; Patent No. 6821754
; GENERAL INFORMATION:
; APPLICANT: Simpson et al.
; TITLE OF INVENTION: Staphylococcus aureus Genes and Polypeptides
; FILE REFERENCE: PB461USDI
; CURRENT APPLICATION NUMBER: US/10/278,946
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: US 09/830,217
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: PCT/US99/06199
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: US 60/078,682
; PRIOR FILING DATE: 1998-03-20

/ REGISTRATION NUMBER: 46,789
/ REFERENCE/DOCKET NUMBER: PB248P1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (240) 314-1224
/ TELEFAX: (301) 309-8439
/ INFORMATION FOR SEQ ID NO: 355:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 668 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ SEQUENCE DESCRIPTION: SEQ ID NO: 355:
US-08-956-171E-355

Alignment Scores: 3.21e-14 Length: 668
Pred. No.: 23.00 Matches: 23
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 6.87% Gaps: 0
DB: 4

US-10-724-972A-6352 (1-335) x US-08-956-171E-355 (1-668)
QY 306 TtpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeu 325
Db 290 TGGGCAAGATCTCGTGGCTTAATTTCTTGAAGAAATGCTTAAGAACTTTGTTGAATTA 349

QY 326 SerLysLys 328
Db 350 TCAAAAAA 358

RESULT 6
US-08-781-986A-355
; Sequence 355, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (301) 309-8504
/ TELEFAX: (301) 309-8512
/ INFORMATION FOR SEQ ID NO: 355:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 668 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
US-08-781-986A-355

QY	169	GlyLysGluGluGluGlyLysSerLyS	177
Db	7073	GGCAAGGAGGAGGAGGAAGAAGA	7047
 RESULT 13			
US-09-949-016-14516/c			
; Sequence 14516, Application US/09949016			
; Patent No. 6812339			
; GENERAL INFORMATION:			
; APPLICANT: VENTER, J. Craig et al.			
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF			
; FILE REFERENCE: CL001307			
; CURRENT APPLICATION NUMBER: US/09/949,016			
; CURRENT FILING DATE: 2000-04-14			
; PRIOR APPLICATION NUMBER: 60/241,755			
; PRIOR FILING DATE: 2000-10-20			
; PRIOR APPLICATION NUMBER: 60/237,768			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR APPLICATION NUMBER: 60/231,498			
; PRIOR FILING DATE: 2000-09-08			
; NUMBER OF SEQ ID NOS: 207012			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 14516			
; LENGTH: 203475			
; TYPE: DNA			
; ORGANISM: Human			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (1)..(203475)			
; OTHER INFORMATION: n = A,T,C or G			
US-09-949-016-14516			
 Alignment Scores:			
Pred. No.:	1.82e+03	Length:	203475
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.69%	Indels:	0
DB:	4	Gaps:	0
 US-10-724-972A-6352 (1-335) x US-09-949-016-14516 (1-203475)			
QY	32	SerRanSerLySGLuserSerLyS	40
Db	67872	AGCAACTCTCCNAGGAGAGCTCTAAG	67846
 RESULT 14			
US-09-949-016-14517/c			
; Sequence 14517, Application US/09949016			
; Patent No. 6812339			
; GENERAL INFORMATION:			
; APPLICANT: VENTER, J. Craig et al.			
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF			
; FILE REFERENCE: CL001307			
; CURRENT APPLICATION NUMBER: US/09/949,016			
; CURRENT FILING DATE: 2000-04-14			
; PRIOR APPLICATION NUMBER: 60/241,755			
; PRIOR FILING DATE: 2000-10-20			
; PRIOR APPLICATION NUMBER: 60/237,768			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR APPLICATION NUMBER: 60/231,498			
; PRIOR FILING DATE: 2000-09-08			
; NUMBER OF SEQ ID NOS: 207012			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 14517			
; LENGTH: 203475			
; TYPE: DNA			
; ORGANISM: Human			
; FEATURE:			
; NAME/KEY: misc_feature			

```
; LOCATION: (1)...(203475)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14517

Alignment Scores:
Pred. No.: 1.82e+03 203475
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.69%
Indels: 0
DB: 4

US-10-724-972A-6352 (1-335) x US-09-949-016-14517 (1-203475)
Qy 32 SerAenSerLySgluSerLyS 40
Db 67872 AGCAACTCTCCAGGAGGAGCTCTTAAG 67846

RESULT 15
US-09-949-016-14518/c
; Sequence 14518, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14518
; LENGTH: 203475
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)...(203475)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14518

Alignment Scores:
Pred. No.: 1.82e+03 203475
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.69%
Indels: 0
DB: 4

US-10-724-972A-6352 (1-335) x US-09-949-016-14519 (1-203475)
Qy 32 SerAenSerLySgluSerLyS 40
Db 67872 AGCAACTCTCCAGGAGGAGCTCTTAAG 67846

RESULT 17
US-09-949-016-17226/c
; Sequence 17226, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17226
; LENGTH: 203475
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)...(203475)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17226

Alignment Scores:
Pred. No.: 1.82e+03 203475
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.69%
Indels: 0
DB: 4

US-10-724-972A-6352 (1-335) x US-09-949-016-17226 (1-203475)
Qy 32 SerAenSerLySgluSerLyS 40
Db 67872 AGCAACTCTCCAGGAGGAGCTCTTAAG 67846
```

```
; LOCATION: (1)...(203475)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14517

Alignment Scores:
Pred. No.: 1.82e+03 203475
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.69%
Indels: 0
DB: 4

US-10-724-972A-6352 (1-335) x US-09-949-016-14517 (1-203475)
Qy 32 SerAenSerLySgluSerLyS 40
Db 67872 AGCAACTCTCCAGGAGGAGCTCTTAAG 67846

RESULT 15
US-09-949-016-14518/c
; Sequence 14518, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14518
; LENGTH: 203475
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)...(203475)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14518

Alignment Scores:
Pred. No.: 1.82e+03 203475
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.69%
Indels: 0
DB: 4

US-10-724-972A-6352 (1-335) x US-09-949-016-14518 (1-203475)
Qy 32 SerAenSerLySgluSerLyS 40
Db 67872 AGCAACTCTCCAGGAGGAGCTCTTAAG 67846

RESULT 16
US-09-949-016-14519/c
; Sequence 14519, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```

Alignment Scores: 1.82e+03 Length: 203475
Pred. No.: 9.00 Matches: 9
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 2.69% Gaps: 0
DB: 4

US-10-724-972A-6352 (1-335) x US-09-949-016-17228 (1-203475)

QY 32 SerAnSerLysGluSerSerLys 40
Db 67872 AGCAACTCTCTCAAGAGAGCTCTAAG 67846

RESULT 20
US-09-949-016-17229/c
; Sequence 17229, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17229
; LENGTH: 203475
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(203475)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17227

Alignment Scores: 1.82e+03 Length: 203475
Pred. No.: 9.00 Matches: 9
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 2.69% Gaps: 0
DB: 4

US-10-724-972A-6352 (1-335) x US-09-949-016-17227 (1-203475)

QY 32 SerAnSerLysGluSerSerLys 40
Db 67872 AGCAACTCTCTCAAGAGAGCTCTAAG 67846

RESULT 19
US-09-949-016-17228/c
; Sequence 17228, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17228
; LENGTH: 203475
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(203475)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17228

Alignment Scores: 1.82e+03 Length: 203475
Pred. No.: 9.00 Matches: 9
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 2.69% Gaps: 0
DB: 4

US-10-724-972A-6352 (1-335) x US-09-949-016-17227 (1-203475)

QY 32 SerAnSerLysGluSerSerLys 40
Db 67872 AGCAACTCTCTCAAGAGAGCTCTAAG 67846

RESULT 18
US-09-949-016-17227/c
; Sequence 17227, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17227
; LENGTH: 203475
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(203475)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17227

Alignment Scores: 1.82e+03 Length: 203475
Pred. No.: 9.00 Matches: 9
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 2.69% Gaps: 0
DB: 4

US-10-724-972A-6352 (1-335) x US-09-949-016-17227 (1-203475)

QY 32 SerAnSerLysGluSerSerLys 40
Db 67872 AGCAACTCTCTCAAGAGAGCTCTAAG 67846

RESULT 19
US-09-949-016-17228/c
; Sequence 17228, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17228
; LENGTH: 203475
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(203475)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17228

```
; SOFTWARE: Patent.pm
; SEQ ID NO 32507
; LENGTH: 181
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 16
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 130
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 131
; OTHER INFORMATION: n=a, g, c or t
; US-09-513-999C-32507

Alignment Scores:
Pred. No.: 23.6 Length: 181
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.39% Indels: 0
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-513-999C-32507 (1-181)
QY 30 SerSerSerAsnSerSerLysGlu 37
Db 163 TCTCTCTCAACTCTTCCAAAGAA 140

RESULT 22
US-09-313-294A-5824
; Sequence 5824, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Ialugudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 5824
; LENGTH: 305
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700350739H1
; US-09-313-294A-5824

Alignment Scores:
Pred. No.: 38.8 Length: 305
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.39% Indels: 0
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-313-294A-5824 (1-305)
QY 306 TrpAlaArgSerArgGlyLeuIle 313
Db 243 TGGCGCGGTTCAAGAGGCTTATC 266

RESULT 23
US-09-513-999C-2950
; Sequence 2950, Application US/09513999C
; Patent No. 6783961
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 2950
; LENGTH: 352
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 111...350
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 119
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 121
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 189
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 260
; OTHER INFORMATION: v=a or c or g
; NAME/KEY: UNSURE
; LOCATION: 4
; OTHER INFORMATION: Xaa=Arg or Thr
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 27
; OTHER INFORMATION: Xaa=Asp or His
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 50
; OTHER INFORMATION: Xaa=Asp or Glu
; US-09-513-999C-2950

Alignment Scores:
Pred. No.: 44.5 Length: 352
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.39% Indels: 0
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-513-999C-2950 (1-352)
QY 28 AsnAsnSerSerSerSerSer 35
Db 148 AACACACAGCAGCAGCAACAGCAGC 171

RESULT 24
US-09-513-999C-29145
; Sequence 29145, Application US/09513999C
; Patent No. 6783961
```

GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 2334
LENGTH: 564
TYPE: DNA
ORGANISM: Myxococcus xanthus
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(564)
OTHER INFORMATION: unsure at all n locations
US-09-902-540-2334

Alignment Scores:
Pred. No.: 69.8 Length: 564
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.39% Indels: 0
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-902-540-2334 (1-564)

Qy 306 TTPAlaArgSerArgGlyLeuIle 313
Db 502 TGGCGCGCTCACGGCGCTCATC 525

RESULT 27
US-09-949-016-39108
Sequence 39108, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 39108
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-39108

Alignment Scores:
Pred. No.: 74.2 Length: 601
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.39% Indels: 0
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-949-016-39108 (1-601)

Qy 171 GluGluGlyLysLysArgLeu 178

GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Duclert, A.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
FILE REFERENCE: 59.US2.REG
CURRENT APPLICATION NUMBER: US/09/513,999C
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent.pm
SEQ ID NO 29145
LENGTH: 378
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 271
OTHER INFORMATION: m=a or c
US-09-513-999C-29145

Alignment Scores:
Pred. No.: 47.6 Length: 378
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.39% Indels: 0
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-513-999C-29145 (1-378)

Qy 93 LysProLeuArgAspLysIleGly 100
Db 178 AAGCGCTAGAGATAAATAGGA 201

RESULT 25
US-09-328-352-884
Sequence 884, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Bratton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 884
LENGTH: 390
TYPE: DNA
ORGANISM: Acinetobacter baumannii
US-09-328-352-884

Alignment Scores:
Pred. No.: 49.1 Length: 390
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.39% Indels: 0
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-328-352-884 (1-390)

Qy 238 LysGlyLeuSerLysTyrLeuLys 245
Db 349 AAAGGCTTGTCTAAATATTAAAA 372

RESULT 26
US-09-902-540-2334
Sequence 2334, Application US/09902540
Patent No. 6833447

Db 69 GAAGAGAGGGGAAAAAAGACTC 92

RESULT 28

US-09-949-016-60203/c

; Sequence 60203, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 60203

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-60203

Alignment Scores:

Pred. No.: 74.2 Length: 601

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.39% Indels: 0

DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-949-016-60203 (1-601)

QY 17 LeuPheValIeuIleAlaThrAla 24

|||||

Db 145 CTCCTTGCTCATTGCTACTGCC 122

RESULT 29

US-09-949-016-60204/c

; Sequence 60204, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 60204

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-60204

Alignment Scores:

Pred. No.: 74.2 Length: 601

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.39% Indels: 0

DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-949-016-60204 (1-601)

QY 17 LeuPheValIeuIleAlaThrAla 24

|||||

Db 461 CTCCTTGCTCATTGCTACTGCC 438

RESULT 30

US-09-949-016-135038

; Sequence 135038, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 135038

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-135038

Alignment Scores:

Pred. No.: 74.2 Length: 601

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.39% Indels: 0

DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-949-016-135038 (1-601)

QY 171 GluGluGluGlyLysArgLeu 178

|||||

Db 69 GAAGAGAGGGGAAAAAAGACTC 92

RESULT 31

US-09-949-016-149234

; Sequence 149234, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 149234

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-149234

Alignment Scores:

Pred. No.: 74.2 Length: 601

Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.39% Indels: 0
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-949-016-149234 (1-601)

QY 204 ValAlaAlaLysSerGlyLeuLeu 211
DB 503 GTGGCTGTCGCAAAAGTGGGCTCTC 526

RESULT 32
US-09-134-001C-909
; Sequence 909, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 909
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-909

Alignment Scores:
Pred. No.: 96.2 Length: 789
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.39% Indels: 0
DB: 3 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-134-001C-909 (1-789)

QY 24 AlaAlaCysGlyAsnAsnSerSer 31
DB 46 GCAGCCTGTGGCAACAATTCATCT 69

RESULT 33
US-09-710-279-1567
; Sequence 1567, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1567
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-1567

Alignment Scores:
Pred. No.: 96.2 Length: 789

Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.39% Indels: 0
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-710-279-1567 (1-789)

QY 24 AlaAlaCysGlyAsnAsnSerSer 31
DB 46 GCAGCCTGTGGCAACAATTCATCT 69

RESULT 34
US-09-543-681A-1850
; Sequence 1850, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 1850
; LENGTH: 843
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-1850

Alignment Scores:
Pred. No.: 103 Length: 843
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.39% Indels: 0
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-543-681A-1850 (1-843)

QY 119 LeuLysProAspLeuIleIleAla 126
DB 256 TTAACCTGCTGATTTAATTGCA 279

RESULT 35
US-09-270-767-14140
; Sequence 14140, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14140
; LENGTH: 886
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-14140

Alignment Scores:
Pred. No.: 108 Length: 886
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.39% Indels: 0
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-270-767-14140 (1-886)

```

QY 28 AsnSerSerSerSerSerSer 35
| | | | | | | | | | | | | | | |
Db 726 ACAAATAGCAGCAGCAACAGC 749

RESULT 36
US-09-489-039A-2246
; Sequence 2246, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2246
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2246

Alignment Scores:
Pred. No.: 119 Length: 981
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.39% Indels: 0
DB: 4 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-489-039A-2246 (1-981)
QY 203 AlaValAlaAlaLysSerGlyLeu 210
| | | | | | | | | | | | | | | |
Db 678 GCCGTGGCTGCANAAAGCGACTG 701

RESULT 37
US-09-641-638-192
; Sequence 192, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET.051CP1
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 192
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 12-515-394 : polymorphic base A or T
; NAME/KEY: misc binding
; LOCATION: 481..500

```

```

; OTHER INFORMATION: 12-515-394.mis1, potential
; NAME/KEY: misc binding
; LOCATION: 502..521
; OTHER INFORMATION: 12-515-394.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 107..126
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 537..557
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc binding
; LOCATION: 489..513
; OTHER INFORMATION: 12-515-394 potential probe
US-09-641-638-192

Alignment Scores:
Pred. No.: 121 Length: 1001
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.39% Indels: 0
DB: 3 Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-641-638-192 (1-1001)
QY 29 AsnSerSerSerSerSerSerLys 36
| | | | | | | | | | | | | | | |
Db 923 AATTCTAGCTCTACAGCTCTATAA 946

RESULT 38
US-09-641-638-193
; Sequence 193, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET.051CP1
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 193
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 12-516-97 : polymorphic base C or T
; NAME/KEY: misc binding
; LOCATION: 481..500
; OTHER INFORMATION: 12-516-97.mis1, potential
; NAME/KEY: misc binding
; LOCATION: 502..520
; OTHER INFORMATION: 12-516-97.mis2, complement
; NAME/KEY: primer_bind
; LOCATION: 405..424
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 891..911
; OTHER INFORMATION: downstream amplification primer, complement

```

US-10-170-097-192

Alignment Scores:

Score:	121	Length:	1001
Percent Similarity:	8.00	Matches:	8
Best Local Similarity:	100.00%	Conservative:	0
Query Match:	2.39%	Mismatches:	0
DB:	4	Indels:	0
		Gaps:	0

US-10-724-972a-6352 (1-335) x US-10-170-097-192 (1-1001)

QY 29 AnSerSerSerAnSerSerLys 36
|||||

Db 923 AATTCTAGCTTAACAGCTCTAAA 946

RESULT 40

US-10-170-097-193
; Sequence 193, Application US/10170097
; Patent No. 6794143
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GEN-T114XC2D1
; CURRENT APPLICATION NUMBER: US/10/170,097
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 193
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 12-516-97 : polymorphic base C or T
; FEATURE:
; NAME/KEY: misc binding
; LOCATION: 481..500
; OTHER INFORMATION: 12-516-97.mis1, potential
; FEATURE:
; NAME/KEY: misc binding
; LOCATION: 502..520
; OTHER INFORMATION: 12-516-97.mis2, complement
; FEATURE:
; NAME/KEY: primer bind
; LOCATION: 405..424
; OTHER INFORMATION: upstream amplification primer
; FEATURE:
; NAME/KEY: primer bind
; LOCATION: 891..911
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc binding
; LOCATION: 489..513
; OTHER INFORMATION: 12-516-97 potential probe
US-10-170-097-193

Alignment Scores:

US-10-724-972a-6352 (1-335) x US-09-641-638-193 (1-1001)

QY 29 AnSerSerSerAnSerSerLys 36
|||||

Db 717 AATTCTAGCTTAACAGCTCTAAA 740

RESULT 39

US-10-170-097-192
; Sequence 192, Application US/10170097
; Patent No. 6794143
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GEN-T114XC2D1
; CURRENT APPLICATION NUMBER: US/10/170,097
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 192
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 12-515-394 : polymorphic base A or T
; FEATURE:
; NAME/KEY: misc binding
; LOCATION: 481..500
; OTHER INFORMATION: 12-515-394.mis1, potential
; FEATURE:
; NAME/KEY: misc binding
; LOCATION: 502..521
; OTHER INFORMATION: 12-515-394.mis2, potential complement
; FEATURE:
; NAME/KEY: primer bind
; LOCATION: 107..126
; OTHER INFORMATION: upstream amplification primer
; FEATURE:
; NAME/KEY: primer bind
; LOCATION: 537..557
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc binding
; LOCATION: 489..513
; OTHER INFORMATION: 12-515-394 potential probe
US-10-724-972a-6352 (1-335) x US-09-641-638-193 (1-1001)

Alignment Scores:

Score:	121	Length:	1001
Percent Similarity:	8.00	Matches:	8
Best Local Similarity:	100.00%	Conservative:	0
Query Match:	2.39%	Mismatches:	0
DB:	3	Indels:	0
		Gaps:	0

Pred. No.: 121
 Score: 8.00
 Length: 1001
 Percent Similarity: 100.00%
 Matches: 8
 Best Local Similarity: 100.00%
 Conservative: 0
 Query Match: 2.39%
 Mismatches: 0
 DB: 4
 Indels: 0
 Gaps: 0

US-10-724-972A-6352 (1-335) x US-10-170-097-193 (1-1001)

Qy 29 AsnSerSerAenSerSerLys 36
 Db 717 AATTCTAGCTCTAACAGCTCTAAA 740

Search completed: November 10, 2005, 03:07:59
 Job time : 664 secs

This page Blank (uspdc)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

QM protein - nucleic search, using frame_plus_p2n model

Run on: November 9, 2005, 19:43:37 ; Search time 5171 Seconds
(without alignments)
3139.141 Million cell updates/sec

Title: US-10-724-972A-6352
Perfect score: 1690
Sequence: 1 GVESVRLKILSVIGLLFVL.....EEMAKELVELSKDKKDK 335

Scoring table:
BLOSUM62 Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cn2.1/USPTO.spool.h/US10724972/runat_03112005.141859.29146/app_query.fasta_1.519
-DB=GenEmbl -QFMT=fastp -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10724972 @CN 1 1 3731 @runat_03112005.141859.29146 -NCPU=3
-NO.MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sv.*
- 13: gb.un.*
- 14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
C 1	1690	100.0 300698 1	AE016750 Staphylococcus epidermidis ATCC 12228
C 2	1233	73.0 307750 1	AP003136 Staphylococcus epidermidis ATCC 12228
C 3	1233	73.0 347650 1	AP003364 Staphylococcus epidermidis ATCC 12228
C 4	1231	72.8 110000 1	Continuation (24 o

ALIGNMENTS

RESULT 1
AE016750/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

AE016750 300698 bp DNA linear BCT 01-JAN-2003
Staphylococcus epidermidis ATCC 12228, section 7 of 9 of the complete genome.
AE016750 AE015929
AE016750.1 GI:27316220
Staphylococcus epidermidis ATCC 12228
Staphylococcus epidermidis ATCC 12228
Bacteria; Firmicutes; Bacillales; Staphylococcus.
1 (bases 1 to 300698)
Zhang, Y., Ren, S., Li, H., Fu, G., Lu, L., Jia, J., Tu, Y., Qin, Z., Chen, Z. and Wen, Y.
Direct Submission
Submitted (05-NOV-2002) Chinese National Human Genome Center at Shanghai, 250 Bi Bo Road, Shanghai 201203, China
Location/Qualifiers
1. 300698
/organism="Staphylococcus epidermidis ATCC 12228"
/mol_type="genomic DNA"
/strain="ATCC 12228"
/db_xref="ATCC:12228"
/db_xref="taxon:176280"
/complement(114..186)
/product="tRNA-Lys"
/complement(190..264)
/product="tRNA-Gln"

Continuation (23 o
AP004829 Staphyloc
AX622368 Sequence
Continuation (11 o
AP004825 Staphyloc
AX583924 Sequence
AX621226 Sequence
AP003132 Staphyloc
AP003361 Staphyloc
A64853 Sequence 10
Continuation (11 o
AX51891 Sequence
D86417 Bacillus su
299108 Bacillus su
Y14084 Bacillus su
299109 Bacillus su
Continuation (12 o
Continuation (12 o
AE006047 Pasteurel
AX433651 Sequence
AB082122 Vibrio pa
AP005089 Vibrio pa
Continuation (12 o
Continuation (13 o
AX933596 Sequence
BX571873 Photorhab
AR354237 Sequence
M26397 K-12 faca g
AF326777 Shigella
AX370197 Sequence
Continuation (46 o
AE016858 Pseudomon
U14003 Escherichia
Continuation (34 o
Continuation (34 o
AX583941 Sequence
BD237137 Staphyloc
AR282355 Sequence
AF079518 Staphyloc
AR354120 Sequence

tRNA	complement (269..352)	/locus_tag="SE1759"
tRNA	/product="tRNA-Tyr"	/codon_start=1
tRNA	complement (369..444)	/transl_table=11
tRNA	/product="tRNA-Val"	/product="multidrug transporter"
tRNA	complement (451..522)	/protein_id="AAO05400.1"
tRNA	/product="tRNA-Glu"	/db_xref="GI:27316225"
tRNA	complement (524..598)	/translation="MNLKSIITVMALILIMFMAAIIETSIISLALPTIKNSLGNLVS
rRNA	/product="tRNA-Asn"	LVTYVFTALVIANPIVGLMSRFKIIYIAVGVLLFALGSLMSGLSQFTFLISRT
rRNA	complement (612..726)	VQGFAGVNMISQIVPKLAFPIPLRYKIMGIVGVSGISSIIGPLGAILLEFASWH
rRNA	complement (886..3713)	WLFYINIPAIVAIIILVLMTPFPDETOVQQRFDIKGLIIFIIYFIALMFGLLNQHH
gene	/product="23S ribosomal RNA"	IIFNFSIILALAVLMLPKIENSIOEPFLTPKEFNISIVLVIITDILLAIITLMGYNL
CDS	complement (2718..2831)	YIPVLOEKLSPLSQSGFVPIPLSVANITLAFNLAKIEAHRKTRKTLTYICSPFVLVS
	/locus_tag="SE1755"	SLMIFGLKPLLIAPAVVAFGLSGFYIYTKDSIVVQESTSPKMKMKMSFFVALTKNL
	/locus_tag="SE1755"	GSSVGSTMGYNALNVGLFGLNHLVGLIIVAVCLLIVMMMLTKSNTIQS"
	/codon_start=1	complement (11146..11829)
	/transl_table=11	/locus_tag="SE1760"
	/product="hypothetical protein"	/locus_tag="SE1760"
	/protein_id="AAO05396.1"	/codon_start=1
	/db_xref="GI:27316221"	/transl_table=11
	/translation="MMIIGGRALFGRGAPLGFLPNSDKLRMPINLTWSEHG"	/product="hemolysin III"
rRNA	complement (3977..5530)	/protein_id="AAO05401.1"
gene	/product="16S ribosomal RNA"	/db_xref="GI:27316226"
CDS	complement (5985..7052)	/translation="MSQSKRKNDSVVEFKDIIPLTFGEIGNAASHGAAALLTLFI
	/locus_tag="SE1756"	LPYAAVHSFNNGTLESISVSIVISIFPMFISSTIYHSNKNNTSHKYLIRIIDHSMI
	/locus_tag="SE1756"	YVAISGTYTPVLSVSGWGLMVLITLLMGTLLWGLIKSLVATKKNHRLSLIIVLVMG
	/codon_start=1	WVGIIPLPIIIMRTSWNFIFPFLGGLSYTIGAFYAQKNKPYFHMIWHIIFVLASFL
	/transl_table=11	HMGIIGFVEM"
	/product="Arp-binding Mrp-like protein"	complement (11848..13035)
	/protein_id="AAO05397.1"	/locus_tag="SE1761"
	/db_xref="GI:27316222"	complement (11848..13035)
	/translation="MLTVNQKEIVGALKDPIIDVPIRESSEGVVDVSIKONINHSVK	/locus_tag="SE1761"
	VAHQLOPOLDIQMAIVKALKENGANTVIGRIEELPSEVVERIYIGSKSPKTIIE	/codon_start=1
	VLMSQNPVEFISLASKGGVCKSTAVNLAVALAREGKVGVLVDADYGFSPQMMGI	/transl_table=11
	DERGIDGKEIIPVERHGKVVISMAFVEENAPVIRGPKMLKMLTNFETSVOMGLD	/product="UDP-N-acetylglucosamine pyrophosphorylase"
	YLLDLPGTDGVALDVHSLPSSKEIIVTPHTAPVAAAGNMAKXHTETILGVI	/protein_id="AAO05402.1"
	ENMSYKESKTEGKEYVFGKGGKKLSDELTQLFAELPQPTNNPNDPSPSIYQSD	/db_xref="GI:27316227"
	DRGELYTLIAKVIASQKQ"	/translation="MLDKNQLEKYNQBEHLEYEKLMSNEKNALDEKVDQLNLAIEIQD
gene	complement (7203..8642)	LYQDLYNNRTIDDSVSEVSEVKSRLENEERHTYEQGYEARNGEFAVLLMAGG
CDS	complement (7203..8642)	QGTGLYKPGKSGPEIETSLFELQARQLRLKEETGHTINWYIMTSDINHKDITFYF
	/locus_tag="SE1757"	KQHKYFNVDANHIHFVKODNIVALSEEGKLVLRDGHIMETPNNGNGVFKSLKKAGYL
	/locus_tag="SE1757"	DKQDQHVKYIPLNNIDNVLVKLDPLPAGFTVTQSKDITQSKITQPKDSSEVGRLVNV
	/codon_start=1	DKQDVTLEYSELDIVNQNNANIGHAFKLGFTSAVDRELPYHLALIKOLKLDEN
	/transl_table=11	FGVVERPTLKFELFYDFRYGTSFVTLQVPREBEFSPLEKNGKSDSVHTATEDLKR
	/product="multidrug resistance protein"	DLI"
	/protein_id="AAO05398.1"	complement (13090..13599)
	/db_xref="GI:27316223"	/locus_tag="SE1762"
	/translation="WASVELSKTKRNLIQSVMLISAFVAIILNQTLLNTALPHIMEELN	complement (13090..13599)
	INENTSQLITGMLVNGVMIPLTAYLMDKVKTRPLYLGMGSLIGISVAIAIAPNGF	/locus_tag="SE1762"
	VLMIAVQVAGAGILMPLMQTLFLFSKEHGRFAMGLAVIQFAPAGIPTFTGLI	/codon_start=1
	IDNASWRVPFIVIGVIALVTFIFGFVSISSYNTKETKIDKRSVIYTLGFLMLYAF	/transl_table=11
	SSAGNLGRSNFVLCSVLISLIIIGIFVKROIITSNPLNLKIFNNKIFCFSTIITSMI	/product="conserved hypothetical protein"
	IMLSWGPALLIPLVQNALGUSALLSLGVIMPGALINGIMSVFTGKFDYKGRPLI	/protein_id="AAO05403.1"
	FTGFTLLISCTFLCFLKVDTSYLYIIRMFPAVSLMLMPPINTTGINALKTDIIS	/db_xref="GI:27316228"
	HGTATMNFGRVMAISGLTAMVTFMSIGAQWVSSSEHASKEMIQRSVAVGVDSFA	/translation="MTGKTHASCGFLVCAITTOYFHTDITTSISVILSVISSILPDI
	LVTVFVITAFVLAFLFKEPNHLTHNSRV"	CHTSQIKGRRLFTSFVRILFGHRTFTHSLFLFIIGISFLFYIQTPTMYMVAIVIGM
gene	complement (8912..9382)	FSHVILDLITLPRGVKLLYPLFPFNIVSPHFHFKTGLGVLDVSLATALSVGAIYTLFQPLN
CDS	/locus_tag="SE1758"	TMMHWLJK"
	/locus_tag="SE1758"	13776..14036
	/codon_start=1	/locus_tag="SE1763"
	/transl_table=11	13776..14036
	/product="conserved hypothetical protein"	/locus_tag="SE1763"
	/protein_id="AAO05399.1"	/codon_start=1
	/db_xref="GI:27316224"	/transl_table=11
	/translation="MMKFLNKSXHLVLTVLTIIVISGATFELTFLGFLGLSLRIL	/product="conserved hypothetical protein"
	LISYIVTTMFIYVHIYVHNHFWMIIIGFLVCKVEYFESKNLNKK"	/protein_id="AAO05404.1"
gene	complement (9570..10907)	/db_xref="GI:27316229"
CDS	/locus_tag="SE1759"	/translation="MAMTVKNDNEVRIOQRVADIKIPNNEIKNVTQDQDIHAVPEEN
	complement (9570..10907)	GKEISRIGSTFGKTRVLDITDQHLIYIYTQNDQKQYNELTK"
		14189..14503
		/locus_tag="SE1764"
		14189..14503
		/locus_tag="SE1764"

```
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA005405.1"
/db_xref="GI:27316230"
/transl_table=11
/translation="MFVVTNRITVKKGAKQMANFTKGGPIBSLKGFEQVWQIDK
DDYSEDYVNSWTEDEDFKNWVNSDVFKQAHKNTGKSDSPVKSIVKSNVLSLN
RR"
gene
complement(14682..16055)
/locus_tag="SE1765"
CDS
complement(14682..16055)
/locus_tag="SE1765"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA005406.1"
/db_xref="GI:27316231"
/transl_table=11
/translation="MNGPKVDIVRGRIKFTMSLGIILFLVPIPVQDGKQOTLP
IAPLAGLKDNLGIMPIILVITIVSGIILCTIYKNKLNPOGLMSSAFNFKIGW
LLELVAFIPFISWTLNIGPEMIKSDTGLVFSLLPTLVAFLFAAIFPLLMYGG
LVBELGPIFRPFLPTFRSTVDNLASFIGDGTGVLTISQYEGEYSRREAV
ISTFSVSTIATVIAETIRMODQFFYFLTVVISCLIAAMTPIKNTPDYVA
KEYSEARNEQLPEGKALKYGFDLATEVIGIKSPKEFLISGFKTVDMFWILFPV
MSGTITATIANVTVPFEIIGKFPVPLELLOIPEAHEASQTILIGFADMFDPSILIE
GQNDVTRFVIGALSISQVLSEVGVILGSKIPVSIKLFMIFLIRTIITPLIIL
LAHLFIG"
gene
complement(16352..17317)
/locus_tag="SE1766"
Alignment Scores:
Pred. No.: 2,67e-115 Length: 300698
Score: 1690.00 Matches: 335
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0
US-10-724-972A-6352 (1-335) x AE016750 (1-300698)
Qy 1 GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuPheValLeu 20
Db 19364 GGAGTGAATCAGTGAGAGGTTTAAANATTTTAAGTGTAAATGGCTATTGTTGTTTAA 19305
Qy 21 IleAlaThrAlaAlaCysGlyAsnAsnSerSerAsnSerSerLysGluSerLys 40
Db 19304 ATTGCAACTGCAGCATGTGGAATAATAGTTCAAGTAACCTCAAGTAAGAGTCATCAAAA 19245
Qy 41 AspGlyValGluIleLysHisGluGluGlyThrLysValProLysHisProLysArg 60
Db 19244 GATGGAGTTGAAATCAAGCAGCAAGAGGCTACTAGCAAGTACCTAAACACCTAAAGCT 19185
Qy 61 ValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProVal 80
Db 19184 GTTGTGTTCTGAGTATTCATTTGTTGATGCTTAGTTGCTTTAGATGTTTAAACCTGTT 19125
Qy 81 GlyIleAlaAspAspAsnLysLysAsnArgIleLysProLeuArgAspLysIleGly 100
Db 19124 GCGATAGCGATGATACAAAAAATTCGTATTATTAAACCATTAAGAGATAAAATTTGGA 19065
Qy 101 LysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLysLys 120
Db 19064 AAATACACTTCTGTAGAACACGTAAGCAACCTAACTTAGAAGAAATCAGTAAACTTAA 19005
Qy 121 ProAspLeuIleIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLys 140
Db 19004 CCAGATTAAATATTGCTGATAATAATAGACACAAAGGTATTTATAAGACTTAAATAAA 18945
Qy 141 IleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAla 160
Db 18944 ATTGCTCTACGATTGAATCAAGTTCAGTGGAGATTATATGAATAATTTATGATGCT 18885
Qy 161 PheLysThrIleSerLysAlaLeuGlyLysGluGluGlyLysArgLeuGluGlu 180
|||||
```

gene	/sub_species="aureus N315"		CDS	complement (4320..4781)		
	/db_xref="taxon:158879"			/gene="SA1857"		
	297..1907			/note="ORFID:SA1857"		
	/gene="SA1853"			/codon_start=1		
	297..1907			/transl_table=11		
	/gene="SA1853"			/product="conserved hypothetical protein"		
	/note="ORFID:SA1853			/protein_id="BAB43139.1"		
	hypothetical protein, similar to DNA mismatch repair			/db_xref="GI:13701847"		
	protein MutS			/translation="MIKINNLDMMQNFAMFLVEQLKSGDLILLNGDLGAGKTTLTQFI		
	/codon_start=1			GKALGVRRITNSPTFNIIKSYRGNLKLHHMCDYCLEDEDELDGFDEFDAQITVIE		
gene	/transl_table=11		CDS	WSQFIKDLLPATHLSINISTISENTRQIELFAQGHEHYEQIEAIIHEFAAH"		
	/protein_id="BAB43135.1"			5292..6980		
	/db_xref="GI:13701843"			/gene="ilvD"		
	/translation="MSTNQFLIFVIAILLTSVIGIVGRYSRQRLKSMETLWQTI			5292..6980		
	SPLEAFIRPNHFQYKLYKEKFESHSLVDKTSWDLNNAIIPHOMNYNLTAIGEMK			/gene="ilvD"		
	LYACLRGMLSTNKSLSLFDNNAEPKRVYTHLALIGKTYTFPPDQITPVRQNIL			/note="ORFID:SA1858"		
	FLCPLPVSFAVFINQSQVIGILLFLMSCLFNIILSATKRTYEDDLKSIIFYASNVILF			/codon_start=1		
	GGYTSIKHAPQEPVNFQKRTARHLTVSLAEVNDEDIGAMVILKVLIFMLDYVLF			/transl_table=11		
	HSTQKSYTHNMLKNCFDIIAELDNHVALMYRRRTLECYTEPQIDTNDGIVFSELT			/product="dihydroxy-acid dehydratase"		
	HPLIADAVANDFSLSQNILTGSNASKSTFKMSIAINIILASAIOTVTASKFYOPG			/protein_id="BAB43140.1"		
gene	IVFTSMANADDVLSGDSYFMABLKSIIKRIVEIPDNQKIYCFIDEIFKGTNTERRIAS		CDS	/db_xref="GI:13701848"		
	ESVLSFHEKSNFRVIAATHDIELAELLKQRYENHPNEVIENNIIHFDYKIKPGKAN			/translation="MRSDMIKKGHQHQPAPRSLLHATGALKSPDTMMKPFVAIINSYID		
	TRNAIELKTSPPAKYVERAKDNVNSG"			IVPGHVLRELADIKAETREAGAIPEFNTIGVDDGIAMGHIGMRYSLPSREIIADA		
	complement (2222..3247)			AETVINAHWFEDGVYIPNCDKITPGMILAMRTNVPDAIFCSGGPMKAGLSAHGKALTL		
	/gene="SA1854"			SSMFEAVGAFEGMSISKEEFLDMEQNACPTCGSCAGMFTANSNCLMEVLGLALPYNG		
	complement (2222..3247)			TALAVSDQREMIROAFAKLVENIKNDLKPDIIVTREAIDDAFALDNAMGSGTNTVLH		
	/gene="SA1854"			TLAINEAGIDYDLERINAIKRTPYLSKTIAPSSSYSMHDVHEAGGVPAIINELMKKDI		
	hypothetical protein, similar to O-sialoglycoprotein			GTLHPDRITVTGTLRENNEGKEIKNFDVIHPDADPYDAQGLSILFGNIAPKGLAP		
	endopeptidase"			VGVDPSPKFTGTGAICFNSHDEAVEAIDNRTVRAGHVIVRYVEGPGKGPMPBLAP		
	/transl_table=11			TSSIVRGKGLDVALITDGRFSGATGIAVGHISPEAASGGPITALIEDGDEITIDLTN		
gene	/codon_start=1		CDS	RTLNVPQEDVLARRRESLTFPKAKVTGYLARYTALVTSANTGGVMQVPENLI"		
	/protein_id="BAB43136.1"			7008..8777		
	/db_xref="GI:13701844"			/gene="ilvB"		
	GVPVEASRHHVEGTTTINEALVDADVSMEDIAIAVTEGPGIGLIGVNAKAL			7008..8777		
	AFAYDKPLIPVHHIAGHYANHIEEPLTFPIALIVSGHTVELVMKDHLSFEVIGET			/gene="ilvB"		
	RDDAVGEAYDKVARTIGLNPQGVDRLAAGEDTYSPPRWLMDKSDYSPFGLKS			/note="ORFID:SA1859"		
	AVINQLNQRQKPIIEANVATSPONSVEVLTPKAIOACKESVQRLIVAGGVASN			/codon_start=1		
	KGLRQSLADQCKVNDIQTIPSPKCLCTDNAAMIGVAGSHLSYQQGRFADALNHSNID			/transl_table=11		
	LEEYSAAE"			/product="acetolactate synthase large subunit"		
	complement (3240..3704)			/protein_id="BAB43141.1"		
gene	/gene="SA1855"		CDS	/db_xref="GI:13701849"		
	complement (3240..3704)			/translation="MSKTOHEVNQNIIDPLKMAELEELEQLNEKTLNDMRSGSEVLVEA		
	/gene="SA1855"			LLKENVDYFGPGGAVLPLDYDTFDGKIHLIARHSGQVAAAEAGVARSGKTVGV		
	hypothetical protein, similar to ribosomal-protein-alanine			VTSRGATNVMGTITDAHCDLPLVFTFGVATPGIGKADAFQEAADILSMTSPITQNY		
	N-acetyltransferase"			QVKRVEDIPKI VHEAFVANSGRGPVVIDFPKDMGVLATNVDLCDEINIPGYEVVT		
	/codon_start=1			PENKIDITFISILLKAAKPVVLGAGINQSNQLLTQFVNKHQITPTVTLLGLGAVP		
	/transl_table=11			YEDTLFLGMGMHGSYASNMALTECDLLINLGSRFDDRLSKDAFAPNAKIVHVDI		
	/protein_id="BAB43137.1"			PSEINKVHVDLGIADCKRPLECLNDKNVETIEHSDMKVHCQNNKQKHPFKLGEDQ		
	/db_xref="GI:13701845"			VFCXPOQTIEYIKITNGEALVITDVGQHQAQFYPFKNHGQWVTSGGLTGMGFG		
	gene	/translation="MDQSKQELNIREMTKEDVPQVDFLERRSFNDSSWTIDAFYHEI		CDS	PSSIGAKLANPDKTVVCFVGQFMQNMALLPEYGLDVKILINNGTLGMWKQMO	
EQNNFAKIFVLEFEQQLIGYGLNIVIDQAQITTVADNNGYVGLGOMLLKYGINVA			1SPTEAVTPMVPSPGKSNHEMGL"			
SHTCDVMSLEVRVNNKVAHYENLFGYGGKRNYYGEGEDANVMWVNLND"			8777..9031			
complement (3677..4366)			/gene="SA1860"			
/gene="SA1856"			8777..9031			
complement (3677..4366)			/gene="SA1860"			
/gene="SA1856"			/note="ORFID:SA1860"			
hypothetical protein, similar to glycoprotein			hypothetical protein, similar to acetolactate synthase			
endopeptidase"			small subunit"			
/codon_start=1			/codon_start=1			
gene	/transl_table=11		CDS	/transl_table=11		
	/protein_id="BAB43138.1"			/protein_id="BAB43142.1"		
	/db_xref="GI:13701846"			/db_xref="GI:13701850"		
	/translation="MKNLRQLSMNSLLIDTSNQPLSVALMONDEVLAETDTSKONH			/translation="NTRILKQVADQVSTLNRTISAFVRLQYNIDTLVHTHSEQPGI		
	SVQLMPAISQLFEQSKAKQDLIAIIVAEQSGSYGLRIGVTVAKTLAYDLVDKLYGV			9168..10172		
	SKLALAKTIDHTDKLLVPVDFARROAVYTGIFQWQGLVLTILEDQTMSEIALRAFL			/gene="ilvC"		
	EDLNQRPYIYGKDTIKLQDELQGEVIAQLPNASVMYQLIDKPTDVHTFTPKYHKLAEA			9168..10172		
	ERNVINSQKNX"			/gene="ilvC"		
	complement (4320..4781)			/note="ORFID:SA1861"		
	/gene="SA1857"			/codon_start=1		
gene	/product="alpha-keto-beta-hydroxylase"		CDS	/transl_table=11		
	reductoisomerase"			/product="alpha-keto-beta-hydroxylase"		
	reductoisomerase"			/product="alpha-keto-beta-hydroxylase"		
	reductoisomerase"			/product="alpha-keto-beta-hydroxylase"		
	reductoisomerase"			/product="alpha-keto-beta-hydroxylase"		
	reductoisomerase"			/product="alpha-keto-beta-hydroxylase"		
	reductoisomerase"			/product="alpha-keto-beta-hydroxylase"		
	reductoisomerase"			/product="alpha-keto-beta-hydroxylase"		
	reductoisomerase"			/product="alpha-keto-beta-hydroxylase"		
	reductoisomerase"			/product="alpha-keto-beta-hydroxylase"		

/protein id="BAB43143.1"
 /db_xref="GI:13701851"
 /translation="MTTVYDQDVKTALQKKKIIVGYGSGQHAAQNLDKNGYDVV
 IGIRGRSFDKAKGSDGDFVPAEAVQADIMVLDFEIQGVYKNEIPNLEKHA
 LAFAGHNIHFVGPDPADVDVFLVAPKPGHLVRRTFVEGSAVPSLFGIQQDASGOA
 RNIALYAKGIGTAGRAGVIETTFKEETEDLFGEOAVLCGVSKLIQSGRETLVYAGY
 QPELAYFEVLHEMKLIYDLMYEGGMENVRYSISNTAEFGDYVSGPRVITDVKENMKA
 VLTDIQNGNFSNFIEDNKNGFKFYLKREBQHQIEKVGREUREMMPFIKSKSIEK
 "
 gene 10202..11731
 /gene="leuA"
 CDS 10202..11731
 /gene="leuA"
 /notes="ORFID:SA1862"
 /codon_start=1
 /transl_table=11
 /product="2-isopropylmalate synthase"

Alignment Scores:

Pred. No.: 6.27e-81 Length: 307750
 Score: 1233.00 Matches: 238
 Percent Similarity: 84.59% Conservative: 42
 Best Local Similarity: 71.90% Mismatches: 51
 Query Match: 72.96% Indels: 1
 DB: 1 Gaps: 0

US-10-724-972A-6352 (1-335) x AP003136 (1-307750)

QY	1	GlyValGluSerValArgGlyLeuLysLeuSerValIleGlyLeuLeuPheValLeu	20
Db	156662	GGAGTGAACG-ATGAGAGCTTAAAACTTTTAGTATATGGGATTAATAGTTGGCTTA	156604
QY	21	IleAlaThrAlaAlaCysGlyAsnAsnSerSerAsnSerSerLysGluSerLys	40
Db	156603	CTTTTAGTGGACCTTGTGGTAATACGGATAATCAAGTAAAAAAGAAATCATCAACTAAA	156544
QY	41	AspGlyValGluIleLysHisGluGluGlyThrLysValProLysHisProLysArg	60
Db	156543	GATACCTATTTCGGTAAAGATGAATGGTACAGTAAAGTAGTACCTAAAGATGCACAACTG	156484
QY	61	ValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProVal	80
Db	156483	ATCGTGTATTAGTAGTACTCATTTGCGAGATGCATTAGCAGCATTAGACCTTTAACAGTT	156424
QY	81	GlyIleAlaAspAsnLysLysAsnArgIleLysPheProLeuArgAspLysIleGly	100
Db	156423	GGTATTCCTGATGATGGTAAAGAAAAACGATATCAATTAACACCTTAAAGAAAAATTCGG	156364
QY	101	LysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluLysSerLysLeuLys	120
Db	156363	GATTATACCTCTGTAGGTACAGTAAACAGCAGCAAACTTAGAGGAAATTAGTAAATTAATA	156304
QY	121	ProAspLeuIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLys	140
Db	156303	CCGGATTTAATATTCGTATAGCAGTACATTAAGGTATTAAATAAGAAATTAACACAA	156244
QY	141	IleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAla	160
Db	156243	ATTGCACCAACATATCATTAAGAGTTTGTATGAGAGACTACAAACAAATTAATTCG	156184
QY	161	PheLysThrIleSerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGlu	180
Db	156183	TTCAAAACATTCCTTAAGCTTTAATAAAGAAAAAGAGCGAAAAACGCTTCTGCTGAG	156124
QY	181	HisAspLysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysVal	200
Db	156123	CATGATAAATTAATCAATTAAGTATTAAGATGAATTAATTTGATGAATTAACAAAAAGTG	156064
QY	201	LeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGly	220
Db	156063	CTTCAGCAGTAGTTGCTTAAAGCTGTTATTAGCACATCCAACTATTCATATGTTGGA	156004
QY	221	GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu	240

Db	156003	CAATTTTAAACGAACCTAGGATTTAAATAATGCATTAAGTCGATGTAACAAAAGGTTTA	155944
QY	241	SerLysTyrLysLysGlyProTyrLeuGluMetAsnThrGluThrLeuSerGlnValAsn	260
Db	155943	AGTAAATATTGTAAGAGGACCTTACTTACAATTAGACACTGAACATTTAGCTGATTAAT	155884
QY	261	ProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGlu	280
Db	155883	CCAGAGCGTATGATCATATGACAGATCATGCTAAAAAAGATCTCTGCTGAATTCAGAAG	155824
QY	281	LeuGluLysAspProValTyrLysLysLeuAsnAlaValLysAsnGlnArgValAspIle	300
Db	155823	TTACAAGAAGATCAACATGGAAGAAAGTTGAATGTCAGTTAAAAAATATCGCTGGATAT	155764
QY	301	LeuAspArgAspLeuTyrPalaArgSerArgGlyLysLeuSerSerGluGluMetAlaLys	320
Db	155763	GTTGACCGTATGTTGGGCAAGATCTCGTGGCTTAATTTCTTCTGAAGAAATGGCTAAA	155704
QY	321	GluLeuValGluLeuSerLysLysAspSerLys	331
Db	155703	GAACTTGTGAATATCAAAAAAAGAACAAAG	155671

RESULT 3

AP003364/c

LOCUS

DEFINITION

AP003364

Staphylococcus aureus subsp. aureus Mu50 DNA, complete genome,

section 7/9

AP003364 BA000017

VERSION

AP003364.3 GI:46395538

KEYWORDS

SOURCE

ORGANISM

Staphylococcus aureus subsp. aureus Mu50

Bacteria; Firmicutes; Bacillales; Staphylococcus.

REFERENCE

AUTHORS

Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I.,

Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Lian, J., Ito, T., Kanamori, M.,

Matsumaru, H., Maruyama, A., Murakami, H., Hosoyama, A.,

Mizutani, U., Y., Takahashi, N., Sawano, T., Inoue, R., Kaito, C.,

Sekimizu, K., Hirakawa, H., Kuhara, S., Goto, S., Yabuzaki, J.,

Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C.,

Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H. and Hiramatsu, K.

Whole genome sequencing of methicillin-resistant Staphylococcus

aureus

Lancet 357 (9264), 1225-1240 (2001)

JOURNAL

MEDLINE

PUBMED

11418146

REFERENCE

AUTHORS

Ohta, T., Hirakawa, H., Morikawa, K., Maruyama, A., Inose, Y.,

Yamashita, A., Oshima, K., Kuroda, M., Hattori, M., Hiramatsu, K.,

Kuhara, S. and Hayashi, H.

Nucleotide Substitutions in Staphylococcus aureus Strains, Mu50,

Mu3, and N315

DNA RESEARCH 11, 51-56 (2004)

3 (bases 1 to 347650)

Direct Submission

Submitted (28-FEB-2001) Masahira Hattori, Kitasato Institute for

Life Sciences, Kitasato University, Kitasato 1-15-1, Sagamiara,

Kanagawa 228-8555, Japan

(E-mail: hattori@genome.its.kitasato-u.ac.jp, Tel: 81-42-778-8194,

Fax: 81-42-778-8193)

On Apr 16, 2004 this sequence version replaced gi:14247707.

This work was done in collaboration with Toshiko Ohta, Mutsumi

Kanamori, Hideo Hayashi (University of Tsukuba), Hideki Hirakawa,

Satoru Kuhara (Kyushu University), Ikuo Uchiyama (Okazaki National

Research Institutes), Susumu Goto, Minoru Kanehisa (Kyoto

University), Naotake Ogasawara (Nara Institute of Science and

Technology) and supported by the Research for the Future Program of

the Japan Society for the Promotion of Science.

Location/Qualifiers

1. 347650

/organism="Staphylococcus aureus subsp. aureus Mu50"

source

FEATURES

gene /mol_type="genomic DNA"
/strain="Mu50"
/sub_species="aureus Mu50"
/db_xref="taxon:158878"
157. .1443
CDS /locus_tag="SAV1936"
157. .1443
/locus_tag="SAV1936"
/codon_start=1
/transl_table=11
/product="similar to aspartate transaminase protein"
/protein_id="BAB58098.1"
/db_xref="GI:14247709"
/translation="MPLAQLNEQLQOSNATAFTMLSDLGQNMFPKGLTQSABAK
STYNATIGMATKNGQMFASSLDAMFNLDTPDEIPYAPPOGIEELRLDWQKMLRD
NPBLSIDNMSLPTDLSLVEALQSYNKDKVIMILNPNPPTGYTPHKEVTIIDAIA
QVPIPKDGHVYTDLSVEALQSYNKDKVIMILNPNPPTGYTPHKEVTIIDAIA
LADQTKVIAVDVADYGLVEDVYVYOSLFTALSNLNAIILPVLGDGATKEFFAQN
RVGPMTCSDQTTKEVLEAKVGLIRSNISGSLPQTQSAVKHVLKNGKQFDEIEQN
IQTLEKRYEYKEVYADQSHHWQADPNFGYFMAIKVHDVDPPEELRKHLEKYSIG
VIALNATDIRAFCEKEKEDIPHFVDSIAKAIDDLR"
complement (1722. .2042)
/locus_tag="SAV1937"
/locus_tag="SAV1937"
/codon_start=1
/note="Bacteriophage phi Sa 3mu"
/transl_table=11
/product="truncated map-w protein"
/protein_id="BAB58099.1"
/db_xref="GI:14247710"
/translation="MKKINKSNKVPVTINKPENIVSNFVFNASKITINDLSIK
LKSAMANDQGITKHIDGLAERAVYKVPKNGSSKVKVDLKTKEYKDERVFKATDIKKVDI
ELK"
complement (2027. .3457)
/locus_tag="SAV1938"
complement (2027. .3457)
/locus_tag="SAV1938"
/note="Bacteriophage phi Sa 3mu"
/codon_start=1
/transl_table=11
/product="truncated map-w protein"
/protein_id="BAB58100.1"
/db_xref="GI:14247711"
/translation="MKFKSLITTTALGVIASTGANFVNTNEASAAKPLDKSSSTLHH
GHSNTQIPYITVNGTSQNLSSLTGKNQNI SYDKENKVSULVFNRGISDIDLRL
SKQAYTVHFKNGTKRVITDLKSGIYADLINTSDIKAI SVNVDTTKQPKDKAKANQV
PVTITVNGTSQNLSSLTGKNQNI SYDKLEKVSLENRGITDVLRLSKQAYT
VNFNGTKVIDLKSIGYTANLINSDDIKSINIVDTKKHLENKARNYQVPYSINL
GTSNLSNLSFSNPNWNYKNLTSQIKSVLKHDRGISQODLKAKKAYTVYFKNGG
KRILQNSKNYTNANLHVAKDVKRIEITVKTGTAKADRYPTIANGTSTPILSKLX
ISNKQLISYKLNDRKSVLSESGISDLDLKFQAKQAYTVYFKNGKKQVNVNLSKDIP
TPNLFSAKDIDKIDVQYTKSKK"
3803. .4042
/gene="truncated-hlb"
/locus_tag="SAV1939"
3803. .4042
/gene="truncated-hlb"
/locus_tag="SAV1939"
/note="Bacteriophage phi Sa 3mu"
/codon_start=1
/transl_table=11
/product="truncated beta-hemolysin"
/protein_id="BAB58101.2"
/db_xref="GI:46395539"
/translation="MHLFCCKLYKRSDNGKTKTNSLKKVATLALANLLVGLTDN
SAKESKDDDTDLKLVSHNVYMLSTVLYPNWRLT"
4414. .4593
/locus_tag="SAV1940"
4414. .4593
/locus_tag="SAV1940"
/note="Bacteriophage phi Sa 3mu"
/codon_start=1

gene /transl_table=11
/product="hypothetical protein"
/protein_id="BAB58102.1"
/db_xref="GI:14247713"
/translation="MKFKTKCKGNIFKRSQEIINNSESEKGLWIGISMLILLFLLI
LFGITACSSSIHFIN"
4617. .4925
CDS /locus_tag="SAV1941"
4617. .4925
/locus_tag="SAV1941"
/note="Bacteriophage phi Sa 3mu"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB58103.2"
/db_xref="GI:46395540"
/translation="MKNHTNIIINILVIVNSLTHFLTNTSFNNSASDFCFIIGAIF
FLIGIFVAIYGMKRAYIWLNLFTNIPVFLHFCVLLLLKVIKPLFIYEGCVLLFT
"
complement (5216. .5566)
/locus_tag="SAV1942"
complement (5216. .5566)
/locus_tag="SAV1942"
/note="Bacteriophage phi Sa 3mu"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB58104.1"
/db_xref="GI:14247715"
/translation="MKIRKSILAGTIAVLASPLVNLNDKNEAQASTLSPTSNEYQNE
KLANEKSLDDELNVNELATGSLTYKTKIKISGQKAMVALKDFPKMSEAKYQIQ
KIYNEIDEALUKSY"
complement (6077. .6373)
/gene="truncated-lytA"
complement (6077. .6373)
/gene="truncated-lytA"
/locus_tag="SAV1943"
/note="Bacteriophage phi Sa 3mu"
/codon_start=1
/transl_table=11
/product="truncated amidase"
/protein_id="BAB58105.2"
/db_xref="GI:46395541"
/translation="MPAGYTLDRKNVPYKKTGYTYTVANVKNVNRDGYSTNSRITGV
LPNNATIKYDQAYCINGYRWITYIANNQRCYIATGEVDKAGNRISSEFGNFSAL"
6461. .6616
/gene="truncated-lytA"
/locus_tag="SAV1943.1"
6461. .6616
/gene="truncated-lytA"
/locus_tag="SAV1943.1"
/note="Bacteriophage phi Sa 3mu"
/codon_start=1
/transl_table=11
/product="truncated amidase"
/protein_id="BAD16608.1"
/db_xref="GI:46395542"
/translation="MRCSDCNISMLRYMYRVKGGKINDICRISVRLGLHGLTGYIVHSY
YIHLIIT"
complement (7063. .7554)
/gene="sak"
/locus_tag="SAV1944"
complement (7063. .7554)
/gene="sak"
/locus_tag="SAV1944"
/note="Bacteriophage phi Sa 3mu"
/codon_start=1
/transl_table=11
/product="staphylokinase precursor"
/protein_id="BAB58106.1"
/db_xref="GI:14247717"
/translation="MLKRSLLFLTIVLLLLFPSSITNEVSASSSFDKGYKKGDDASY

FEPTGPMVNVTVGDKGNELLSPHVEPIKPGTTLTKIEIYVVEWALDATAYKE
FRVLEDFSAKIEVYYDKNKKKEETKSPFITEKGFVVPDLSEHIKNPGFNLIKVI
EKK"
gene complement (7769. .8500) 347650
/locus.tag="SAV1945"
CDS complement (7769. .8500)

Alignment Scores:

Pred. No.: 7.26e-81 Length: 347650
Score: 1233.00 Matches: 238
Percent Similarity: 84.59% Conservative: 42
Best Local Similarity: 71.90% Mismatches: 51
Query Match: 72.96% Indels: 1
DB: 1 Gaps: 0

US-10-724-972A-6352 (1-335) x AP003364 (1-347650)

QY	1	GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeu 20
DB	239361	GGAGTGAACG-ATGAGAGGTCTAAAAACCTTTTAGTATATTTGGGATTAATAGTTGCTTAA 239303
QY	21	IleAlaThrAlaAlaCysGlyAsnAsnSerSerAsnSerSerLysGluSerLys 40
DB	239302	CTTTTAGTTGCACCTTGTGTAATACGGATAATCAAGTAAAAAAGAAATCATCACTAAA 239243
QY	41	AspGlyValGluIleLysHisGluGluGlyThrThrLysValProLysHisProLysArg 60
DB	239242	GATACATATTTCCGTAAAAAGATGAAAAATGGTACAGTAAAAAGTACCTAAAGATGCAAAACCT 239183
QY	61	ValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProVal 80
DB	239182	ATCGTTGTATTAGAGTACTCATTTTGCAGATGCAATAGAGCAATAGACGTTAAACCGATT 239123
QY	81	GlyIleAlaAspAsnLysLysAsnArgIleLeuLysProLeuArgAspLysIleGly 100
DB	239122	GGTATTGCTGATGATGGTAAGAAAAAACGTATCATTTAAACCCAGTTAGAGAAAAATTTGGG 239063
QY	101	LysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluLysSerLysLeuLys 120
DB	239062	GATTATACCTCTGTGCTACAGTAAACAGCCAAACCTTAGAGGAATTTAGTAAATTTAAA 239003
QY	121	ProAspLeuIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLys 140
DB	239002	CCGGATTAAATTCGCTGATAGCATAGACATAAGAGTATTAAATAAGAAATTTAAACAAA 238943
QY	141	IleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAla 160
DB	238942	ATTGCCACCAACATTATCATTAAGAGTTTTGTAGGAGACTACAAACAAATATTAAATTCG 238883
QY	161	PheLysThrIleSerLysAlaLeuGlyLysGluGluGluGlyLysArgLeuGlu 180
DB	238882	TTCAAAACAATTTGCTAAAGCTTTAAATAAGAAAAAAGAGGCGAAAAACGCTTCTGCTGAG 238823
QY	181	HisAspLysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysVal 200
DB	238822	CATGATAAATTAATCAATAAGTATAAAGATGAATTAATTAATTTGATAGAAATCAAAAAAGTG 238763
QY	201	LeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGly 220
DB	238762	CTTCCAGCAGTAGTTGCTAAAGCTGTTATTAGCACATCCAAACATATTCATATGTTGGA 238703
QY	221	GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAspValThrLysGlyLeu 240
DB	238702	CAATTTTAAACGAACCTAGGATTTAAATAATGCATTAAGTGCAGATGTAAACAAAAGGTTTA 238643
QY	241	SerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsn 260
DB	238642	AGTAAATATTTGAAAGGACCTTACTTACAAATTAGACACTGAAACATTTTAGCTGATTTAAAT 238583
QY	261	ProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGlu 280
DB	238582	CCAGAGCGTATGATCATTTATGACAGATCATGCTAAAAAAGATTTCTGCTGAATTTCAAGAAG 238523

QY	281	LeuGluLysAspProValTTPlyLysLeuAsnAlaValLysAsnGlnArgValAspIle 300
DB	238522	TTACAGAAGATGCAACATGGAATAAGTTGAATGCAGTTAAAAATTAATCCGCTGGATATT 238463
QY	301	LeuAspArgAspLeuTyrPalaArgSerArgGlyLeuIleSerSerGluGluMetAlaLys 320
DB	238462	GTTGACCGTGATGTTTGGGCAAGATCTCGTGGCTTAATTTCTTCTGAGAAATGGCTAAA 238403
QY	321	GluLeuValGluLeuSerLysLysAspSerLys 331
DB	238402	GAACTTGTGTAATTATCAAAAAAAGAACAAAAG 238370

RESULT 4

BX571856_23/c

WPCOMMENT

Sequence split into 29 fragments LOCUS BX571856 Accession BX571856

Fragment Name	Begin	End
BX571856_00	1	110000
BX571856_01	100001	210000
BX571856_02	200001	310000
BX571856_03	300001	410000
BX571856_04	400001	510000
BX571856_05	500001	610000
BX571856_06	600001	710000
BX571856_07	700001	810000
BX571856_08	800001	910000
BX571856_09	900001	1010000
BX571856_10	1000001	1110000
BX571856_11	1100001	1210000
BX571856_12	1200001	1310000
BX571856_13	1300001	1410000
BX571856_14	1400001	1510000
BX571856_15	1500001	1610000
BX571856_16	1600001	1710000
BX571856_17	1700001	1810000
BX571856_18	1800001	1910000
BX571856_19	1900001	2010000
BX571856_20	2000001	2110000
BX571856_21	2100001	2210000
BX571856_22	2200001	2310000
BX571856_23	2300001	2410000
BX571856_24	2400001	2510000
BX571856_25	2500001	2610000
BX571856_26	2600001	2710000
BX571856_27	2700001	2810000
BX571856_28	2800001	2902619

Continuation (24 of 29) of BX571856 from base 2300001 (BX571856 Staphylococcus aureus su

Alignment Scores:

Pred. No.: 2.58e-81 Length: 110000
Score: 1231.00 Matches: 239
Percent Similarity: 83.99% Conservative: 39
Best Local Similarity: 72.21% Mismatches: 53
Query Match: 72.84% Indels: 1
DB: 1 Gaps: 0

US-10-724-972A-6352 (1-335) x BX571856_23 (1-110000)

QY	1	GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeu 20
DB	50332	GGAGTGAACG-ATGAGAGGTCTAAAAACCTTTTAGTATATTTGGGATTAATAGTTGCTTAA 50274
QY	21	IleAlaThrAlaAlaCysGlyAsnAsnSerSerAsnSerSerLysGluSerLys 40
DB	50273	TTTTTAGTTCAGCTTCTGGTAAATACGGATAATTTCAAGTAAAAAAGAAATCATCACTAAA 50214
QY	41	AspGlyValGluIleLysHisGluGluGlyThrThrLysValProLysHisProLysArg 60
DB	50213	GATACATATTTCCGTAAAAAGATGAAAAATGGTACAGTAAAAAGTACCTAAAGATGCAAAAACGT 50154
QY	61	ValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProVal 80
DB	50153	ATCGTTGTATTAGAGTACTCATTTGACAGATGATTTGACAGCATTAGACGTTAAACCGATT 50094

QY	81	GlyVleAlaAaspAenLysLysAenArgIleileysProLeuArgAaspLysIleGly	100
Db	50093	GGTATTGCTGATGATGAAGAAAAACGTTATTAACCGATTAGAGAAAAATTCGG	50034
QY	101	LysTyThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLys	120
Db	50033	GATTATATCTCTGTAGGTACACGTAACACCCAACTTAGAGGAAATTAATTAATA	49974
QY	121	ProAspLeuIleAlaAaspAenAsnArgHileGlyIleTyLysAaspLeuAenLys	140
Db	49973	CCGGATTAAATTCGCTGATAGCAGTACAGTAAAGGATTAAATAAAGAAATTAACA	49914
QY	141	IleAlaProThrIleGluLeuLysSerPheAaspGlyAspTyrsAsnGluAenIleAasp	160
Db	49913	ATTGCACCAACATTCATTAAGAGTTTGTGGTGACATACAGCAAAATATCGATGCT	49854
QY	161	PhelysThrIleSerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGlu	180
Db	49853	TTCAAGACAAATTCGGAAGCTTTAGATAAGAAAAAGAGCGCAAAACGCTTCCTG	49794
QY	181	HieAaspLysIleGluGlyLysGluIleThrMetAaspLysAenGlnLysVal	200
Db	49793	CATGATTAATTAATCAATAAGATATAAGATGAAATTAATTTGATAGAAATCAAAA	49734
QY	201	LeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAenSerTyValGly	220
Db	49733	CTTCAGCAGTTGTTGCAAGCTGTTATAGCACATCCAACTATTCGTATGTTGGA	49674
QY	221	GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAaspValThrLysGlyLeu	240
Db	49673	CAATTTTTAAACGAACCTTGGATTTAAATAATGCATTAAGTATGATGATACAAA	49614
QY	241	SerLysTyThrLeuLysGlyProTyThrLeuGlnMetAenThrGluThrLeuSerGln	260
Db	49613	AGTAATACTTGAAGGACCTTACTTACATAGTACTGAACTTAACTAGCTGACTTA	49554
QY	261	ProGluArgMetPheIleMetThrAsnLysAlaSerSerAenGluProSerLeuLysGlu	280
Db	49553	CCTGAACGCATGATTTATATGACAGATTAATGCTAAAAAGATTCCTGGAATTC	49494
QY	281	LeuGluLysAaspProValTropLysLysLeuAenAlaValLysAenGlnArgValAasp	300
Db	49493	TTACAAGAAAGTGCACAACTGGAAGAAATTTGAATGCGATTAAATAATTCGTG	49434
QY	301	LeuAaspArgAaspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluMetAlalys	320
Db	49433	GTTGACCGTGATGTTTGGCAAGATCTCGTGGCTTAATTTCTCTCAAGAAATG	49374
QY	321	GluLeuValGluLeuSerLysLysAaspSerLys	331
Db	49373	GAACTTTGTAATTCATCAAAAAAGAAACAAAG	49314

RESULT 5
BX571857_22/c
WPCOMMENT

Sequence split into 28 fragments LOCUS BX571857 Accession BX571857

Fragment Name	Begin	End
BX571857_00	1	110000
BX571857_01	100001	210000
BX571857_02	200001	310000
BX571857_03	300001	410000
BX571857_04	400001	510000
BX571857_05	500001	610000
BX571857_06	600001	710000
BX571857_07	700001	810000
BX571857_08	800001	910000
BX571857_09	900001	1010000
BX571857_10	1000001	1110000
BX571857_11	1100001	1210000
BX571857_12	1200001	1310000
BX571857_13	1300001	1410000
BX571857_14	1400001	1510000

BX571857_15	1500001	1610000
BX571857_16	1600001	1710000
BX571857_17	1700001	1810000
BX571857_18	1800001	1910000
BX571857_19	1900001	2010000
BX571857_20	2000001	2110000
BX571857_21	2100001	2210000
BX571857_22	2200001	2310000
BX571857_23	2300001	2410000
BX571857_24	2400001	2510000
BX571857_25	2500001	2610000
BX571857_26	2600001	2710000
BX571857_27	2700001	2799802

Continuation (23 of 28) of BX571857 from base 2200001 (BX571857 Staphylococcus aureus str)

Alignment Scores:	4,34e-81	Length:	110000
Pred. No.:	1228.00	Matches:	237
Score:	84.29%	Conservative:	42
Percent Similarity:	71.60%	Mismatches:	52
Best Local Similarity:	72.66%	Indels:	1
Query Match:	1	Gaps:	0
DB:			

US-10-724-972A-6352 (1-335) x BX571857_22 (1-110000)

QY	1	GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuPheValLeu	20
Db	48090	GGAGTGGAAACG-ATGAGAGCTCTAAAAACCTTTTAGTATATTGGGATTAATAGTTGCCTTA	48032
QY	21	IleAlaThrAlaAlaCysGlyAsnAenSerSerAenSerSerLysGluSerSerLys	40
Db	48031	CTTTTAGTTCAGCTTGTGTAATACGGATAATTCAGTAAAAAAGAAATCATCACTAAA	47972
QY	41	AspGlyValGluIleLysHisGluGluGlyThrThrLysValProLysHisProLysArg	60
Db	47971	GATACCTATTCGGTAAAGATGAAAAATGGTACAGTAAAGTACCTAAAGATGCAAAACGT	47912
QY	61	ValValValLeuGluTyThrSerPheValAspAlaLeuValAlaLeuAaspValLysProVal	80
Db	47911	ATCGTTGTTAGTAGTACTCATTTGCAGATGCACTTAGCAGCATTAGACGTTAAATCCAGTT	47852
QY	81	GlyIleAlaAaspAenLysLysAenArgIleileysProLeuArgAaspLysIleGly	100
Db	47851	GGTATTGCTGATGATGGTAAGAAAAACGTTATCAATTAACCCAGTTTAGAAAAAATTTGGG	47792
QY	101	LysTyThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLys	120
Db	47791	GATTATATCTCTGTAGGTACAGTAAACAGCCAACTTTAGAAGAAATTAGTAAATTAATA	47732
QY	121	ProAspLeuIleAlaAaspAenAsnArgHileLysGlyIleTyLysAaspLeuAenLys	140
Db	47731	CCGGATTTAAATTCGCTGATAGCAGTACATAAAGGTTATTAATAAAGAAATTAACAA	47672
QY	141	IleAlaProThrIleGluLeuLysSerPheAaspGlyAspTyrsAsnGluAenIleAasp	160
Db	47671	ATTGCACCAACATTCATTAAGAGTTTGTAGTACAGTACAAACAAACATTAATTCG	47612
QY	161	PhelysThrIleSerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGlu	180
Db	47611	TTCAAAACAAATTCCTAAAGCTTTAAATAAAGAAAAAGAGCGCAAAACGCTTCGTGAG	47552
QY	181	HieAaspLysIleGluGluGlyLysGluIleThrMetAaspLysAenGlnLysVal	200
Db	47551	CATGATAAATTAATCAATAAGATATAAAGATGAATTAATTTGATAGAAATCAAAAGTG	47492
QY	201	LeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAenSerTyValGly	220
Db	47491	CTTCAGCAGTATGCTTAAGCTGTTTATTAGCACATCCAACTATTCATATGTTGGA	47432
QY	221	GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAaspValThrLysGlyLeu	240
Db	47431	CAATTTTTAAACGAACATGAGATTTAAAAAATGCATTAAAGTACGATGCAAAAGGTTTA	47372


```

/gene
CDS
/protein_id="BAB95818.1"
/db_xref="GI:21205124"
/translation="MYVKQLKFSDEARQAMLRGVQDQLANAVKVTIGKGRNVLDKEFT
LKNVTSANGPVQRQIQIDRAVKVAVALHNSQKVENQNEIAQVGAISAADIEIGRYI
SEAMEKGVNDVLTIEESNGLTNTELEVEGMDQFDRGQYPMVTDSDKVAELERFYI
LVTDKRSQFDQILPLEQVQSNRPILVADEVEGDALTNVLNMRGTFTAVAKA
PGDGRKAMLEDLAULTGAQVITDDGLDKDASIDMLGTASKVEVTDNTVVDGD
LGENSDARVQSQSIETESQFDRKELQKLAGGAVGVKVAASSTELKRGD
RIDALNSTRAAVEGIVAGGVTALNVYQKVEIEABGDIEVGNVILKALPAPVQ
IAENAGLEGSVIVERLKNABPGVGFNAATNEWNMLEAGIVDPTKTVRSALQHAASVA
AMFLTAAVVASIPEKNNDQPNMGPGMM"
/complement(5125..5409)
/gene="groES"
/complement(5125..5409)
/gene="groES"
/notes="ORFID:MM1954"
/codon_start=1
/transl_table=11
/product="GroES protein"
/protein_id="BAB95819.1"
/db_xref="GI:21205125"
/translation="MLKPIGNRVIIIEKKEQETTKSGIVLTDSAKESNKGVI VAGT
GRLLNDGTRVTPVEKGDVRVFOQYAGTEVKRNETVTLVNEEDILAVIE"
5584..6327
/gene="MW1955"
5584..6327
/gene="MW1955"
/notes="ORFID:MM1955"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB95820.1"
/db_xref="GI:21205126"
/translation="WTRLSALLTVIIVLSQFLPLIVKLPFVQVSGIELTKAVIY
IQVLPIAATIIILINKIKNPTKLELVKEPKYIIPWALLGFALVMIQMVSVI
LTQYGGQQSPNTEKLIIRKPIIFPFVSIIGPLLEEYVFKVIFGELFNAIKGN
RIVAFIATTVSSIIIFALAHNDFKPIPVYFGMVGFISLAYVWTKRLAVPIIHLQNG
FWVIFQLNPEALKKATEQANFIHFIP"
/complement(6352..7611)
/gene="MW1956"
/complement(6352..7611)
/gene="MW1956"
/notes="ORFID:MM1956"
/hypothetical protein, similar to SdrH"
/codon_start=1
/transl_table=11
/protein_id="BAB95821.1"
/db_xref="GI:21205127"
/translation="MSYHWFKKMLLSTLILSSSSIGLATHVTEAKONLNGEKPTTN
LNHNITSPVNSEMMNNETGTPHESNQTGNSRDANDPDSNNVKNQNNQPSST
DKPDNNQNSPNPKPDNPKPDNPKPDNPKPDNPKPDNPKPDNPKPDNPKPDN
DKPDNPKPDNPKPDNPKPDNPKPDNPKPDNPKPDNPKPDNPKPDNPKPDN
ADGNSQGWQNGQNSQNTGNDVFSQREBALANGAYKYNPYILNQINKLGKDYG
EYTDIDYNIIRKQNFSGNAYLNGLQQSNYFRFQYFPLPKSERYRNLDEVLALIT
GSIQSPDLKKPEKDPKQSKQSFLEKEDDFTTVVKQEDNKKKSASTAYSKSWLAIVCS
MMVFSIMLFLVKRNKKKNKQESQRR"
7808..8434
/gene="MW1957"
7808..8434
/gene="MW1957"
/notes="ORFID:MM1957"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB95822.1"
/db_xref="GI:21205128"
/translation="MGLFKDKKAMFTDNAMERRSIYNLKDISISDDLESMTAHA
VHVPSFNSQSTRIIVLLNDNNKFWNTKAILKEVMGNDFFETEQKIDNFRHSY
GTLFYEDQDVVSGIQEOMPNYDNEFATWSTQTNAMHQFAIWTALATKIGASLOHYN
PLVDWTSNEENIPKSMKLIQAMPFGDIRAAGEKTFNPVEDRFVIKK"
8795..9580
/gene="MW1958"

```

```

CDS
8795..9580
/gene="MW1958"
/notes="ORFID:MM1958"

Alignment Scores:
Pred. No.: 1,47e-80 Length: 304050
Score: 1228.00 Matches: 237
Percent Similarity: 84.29% Conservative: 42
Best Local Similarity: 71.60% Mismatches: 52
Query Match: 72.66% Indels: 1
DB: 1 Gaps: 0

US-10-724-972A-6352 (1-335) x AP004829 (1-304050)
Qy 1 GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeu 20
Db 171485 GGAGTGGACG-ATGAGAGCTCTAAAAAATTTAGTATATTTAGTTGGCTTA 171427
Qy 21 IleAlaThrAlaAlaCysGlyAsnAsnSerSerSerSerSerSerSerSerSerSer 40
Db 171426 CTTTGTAGTTCAGCTTGTGTAATACGGATAATTCAGATAAAAAAGAAATCAACATAA 171367
Qy 41 AspGlyValGluLeuLysHisGluGlyThrThrLysValProLysHisProLysArg 60
Db 171366 GATACATATTTCCGTTAAAGATGAATAATGGTACAGTAAAAAGTACCTAAAGATGCAAAACGT 171307
Qy 61 ValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProVal 80
Db 171306 ATCGTTGTATTAGAGTACTCATTTTCAGATGCATTCAGCAGATTAGACGTTAATCCAGTT 171247
Qy 81 GlyIleAlaAspAspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGly 100
Db 171246 GGTATTCCTGATGATGTAAGAAAAAAGCTATCATTAACACGATTAGAGAAAAAATGGG 171187
Qy 101 LysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluLysIleSerLysLeuLys 120
Db 171186 GATTATATCTCTGTAGGTACACGTAAACAGCAAACTTAGAAGAAATTAGTAATTTAAA 171127
Qy 121 ProAspLeuIleIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLys 140
Db 171126 CCGGATTTAATATTCGCTGATAGCAGTACATTAAGGTATTAAATAAGAAATTAACACAA 171067
Qy 141 IleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAla 160
Db 171066 ATTGCACCAACATTATCATTAAGAGATTTTCATGAGAGACTACAAACAAACATTAATTCG 171007
Qy 161 PheLysThrIleSerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGlu 180
Db 171006 TTCAAAAACAATTGCTAAAGCTTTAAATAAGAAAAAGGCGGAAAAACGCTTCCTGAG 170947
Qy 181 HisAspLysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysVal 200
Db 170946 CATGATAAATTAATCAATAGTATAAGATGAATTAATTTGTAGAAATCAAAAGTG 170887
Qy 201 LeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGly 220
Db 170886 CTTCCAGCAGTAGTTGCTAAAGCTGTTTATTAGCACATCCAAACTATTTCATATGTGGA 170827
Qy 221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAspValThrLysGlyLeu 240
Db 170826 CAATTTTAAACGAACACTAGGATTTAAAAATGCAATTAAGTGAAGTGTACAAAGGTTTA 170767
Qy 241 SerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsn 260
Db 170766 AGTAATAATTTGAAAGGACCTTACTTACAAATTTAGACACTGAACATTTAGCTGATTAAT 170707
Qy 261 ProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGlu 280
Db 170706 CCAGAGCGTATGATCATTTATGACTGATCAACGCTAAAAAAGATTTCTGCTGAATTCAGAAA 170647
Qy 281 LeuGluLysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIle 300
Db 170646 TTACAGAAGAGTCAACACGGAAGAAAAATTTGAATGCGAGTTAAAAATAATTCGCTGGATATT 170587

```

QY 301 LeuAepArgLeuTTPAlaAerSerArgGlyLeuileSerSerGluGluMetAlaLys 320
Db 170586 GTTACCGTGTATGTTGGGCAAGATCTCGTGGCTTAATTTCTTGAAGAAATGGCTAAA 170527

QY 321 GluLeuValGluLeuSerLysLeuAepSerLys 331
Db 170526 GAACCTGTTGAATATCAAAAAAGAACAAAAG 170494

RESULT 7
AX622368
LOCUS AX622368 1044 bp DNA linear PAT 20-FEB-2003
DEFINITION Sequence 5331 from Patent WO02094868.
ACCESSION AX622368
VERSION AX622368.1 GI:28450503
KEYWORDS
SOURCE Staphylococcus aureus
ORGANISM Staphylococcus aureus
Bacteria; Firmicutes; Bacillales; Staphylococcus.

REFERENCE 1
AUTHORS Masignani, V.C., Mora, M.C. and Scarselli, M.C.
TITLE Staphylococcus aureus proteins and nucleic acids
JOURNAL Patent: WO 02094868-A 5331 28-NOV-2002;
Chiron Spa (IT)

FEATURES
Location/Qualifiers
1..1044
/organism="Staphylococcus aureus"
/mol_type="unassigned DNA"
/db_xref="taxon:1280"

ORIGIN
Alignment Scores:
Pred. No.: 1,83e-82 Length: 1044
Score: 1214.00 Matches: 237
Percent Similarity: 84.59% Conservative: 43
Best Local Similarity: 71.60% Mismatches: 51
Query Match: 71.83% Indels: 2
DB: 6 Gaps: 0

US-10-724-972A-6352 (1-335) x AX622368 (1-1044)
QY 1 GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeu 20
Db 54 GGAGTGGAAAG-ATGAGAGGCTTAAACCTTT-AGTATATTGGGATTAAATAGTTCCTTA 111
QY 21 IleAlaThrAlaAlaCysGlyAenAenSerSerAenSerSerLysGluSerLys 40
Db 112 CTTTATGTTGAGCTTGTGTAATACGGATAATTCAGTAAAGAAATCATCAACTAAA 171
QY 41 AspGlyValGluIleLysHisGluGlyThrLysValProLysHisProLysArg 60
Db 172 GATACATTTTCGTTAAAGATGAAATGTTACAGTAAAGTACCTAAAGATGCAAAAGCT 231
QY 61 ValValValLeuGlyThrSerPheValAlaLeuValAlaLeuAepValLysProVal 80
Db 232 ATCGTTGTATTAGTACTCTATTCGACATGCTAGCAGCATTTAGACGTTAAACAGTT 291
QY 81 GlyIleAlaAepAepAenLysAenArgIleLysProLeuAerGlyLeuGly 100
Db 292 GGTATTCGTGATGATGTTAAGAAAACGATATCAATTAACCGTTAGAGAAAAATGGG 351
QY 101 LysThrThrSerValGlyThrArgLysGlnProAenLeuGluGluIleSerLysLys 120
Db 352 GATTATATCTCTGAGGTACAGTAAACAGCAAACTTAGAAGAAATTTAGTAATATAA 411
QY 121 ProAepLeuIleAlaAepAenAenArgHisLysGlyIleThrLysAepAenLys 140
Db 412 CCGGATTTAATTCGCTGATGACAGTACATAAAGGTATTATAAAGAAATTAACAAA 471
QY 141 IleAlaProThrIleGluLeuLysSerPheAepGlyAepTyrAenGluAenIleAepAla 160
Db 472 ATTCACCAACATTTATCATTAAGAGTTTGTAGAGACTACAAACAAATATTAAATTCG 531

QY 161 PheLysThrIleSerLysAlaLeuGlyLysGluGluGlyLysArgLeuGluGlu 180
Db 532 TTCAAAACATTTGCTAAAGCTTTAAATAAAGAAAAGAGCGGAAAACGCTTCCTGCTAG 591
QY 181 HisAepLysLysIleGluGluThrLysLysGluIleThrMetAepLysAenGlnLysVal 200
Db 592 CATGATAAATTAATCAATAGTATAAAGATGAAATTAATTTGATAGAAATCAAAAGTG 651
QY 201 LeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAenSerTyrValGly 220
Db 652 CTTCCAGCAGTAGTCTTAAAGCTGGTTATTAGCACATCCAAACTATTTCATATGTTGA 711
QY 221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAepValThrLysGlyLeu 240
Db 712 CAAATTTTAAACGACTAGGATTTAAATATGCAATTAAGTACCATGTAACAAAGGTTTA 771
QY 241 SerLysThrLysGlyProLysLeuGlnMetAenThrLysLeuSerGlnValAsn 260
Db 772 AGTAAATATTGAAAGGACCTTACTTACAATTAGACACTGAACATTAGCTGATTAAAT 831
QY 261 ProGluArgMetPheIleMetThrAenLysAlaSerSerAenGluProSerLeuLysGlu 280
Db 832 CCAGAGCGTATGATCATTTATGACAGATCATGCTTAAAGAAAGATTCTGCTGAATTCAGAAG 891
QY 281 LeuGluLysAepProValTrpLysLysLeuAenAlaValLysAenGlnArgValAspIle 300
Db 892 TTACAAGAGATGCAACATGGAAAAGTTGAATGCAATTAATAAATAATCCGCTGGATATT 951
QY 301 LeuAepArgAepLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLys 320
Db 952 GTTGACCGGTGATGTTTGGGCAAGATCTCGTGGCTTAATTTCTTGAAGAAATGGCTAAA 1011
QY 321 GluLeuValGluLeuSerLysLysAepSerLys 331
Db 1012 GAACCTGTTGAATTTCAAAAAAGAACAAAAG 1044

RESULT 8
BX571857.10
WPCOMMENT

Sequence split into 28 fragments LOCUS BX571857 Accession BX571857

Fragment Name	Begin	End
BX571857_00	1	110000
BX571857_01	100001	210000
BX571857_02	200001	310000
BX571857_03	300001	410000
BX571857_04	400001	510000
BX571857_05	500001	610000
BX571857_06	600001	710000
BX571857_07	700001	810000
BX571857_08	800001	910000
BX571857_09	900001	1010000
BX571857_10	1000001	1110000
BX571857_11	1100001	1210000
BX571857_12	1200001	1310000
BX571857_13	1300001	1410000
BX571857_14	1400001	1510000
BX571857_15	1500001	1610000
BX571857_16	1600001	1710000
BX571857_17	1700001	1810000
BX571857_18	1800001	1910000
BX571857_19	1900001	2010000
BX571857_20	2000001	2110000
BX571857_21	2100001	2210000
BX571857_22	2200001	2310000
BX571857_23	2300001	2410000
BX571857_24	2400001	2510000
BX571857_25	2500001	2610000
BX571857_26	2600001	2710000
BX571857_27	2700001	2799802

Continuation (11 of 28) of BX571857 from base 1000001 (BX571857 Staphylococcus aureus st

Alignment Scores:

Pred. No.: 2.61e-43 Length: 110000

Score: 725.50 Matches: 149
Percent Similarity: 66.67% Conservative: 63
Best Local Similarity: 46.86% Mismatches: 95
Query Match: 42.93% Indels: 11
DB: 1 Gaps: 4

US-10-724-972A-6352 (1-335) x BX571857_10 (1-110000)

Qy 13 ValileGlyLeuLeuPheValLeuLeuLeuAlaThrAlaAlaCysGlyAsnAsnSerSerSer 32
Db 42342 GTTGTGTTATGCTATCTAGTGTGACGAGTGGGGTGGTGGTCAAAAGATACT--- 42398
Qy 33 AsnSerSerLysGluSerSerLysAspGlyValGluLeuLeuHisGluGluGlyThr 52
Db 42399 -----GAAGAGAAAACCTGAATGACGACAAATAAAGATGAATAGGAACTGAA 42446
Qy 53 LysValProLysHisProLysArgValValValLeuGluLysSerPheValAspAlaLeu 72
Db 42447 AAAATTAGAAAATCTCTAAACGTGTGTGTATTAGATATAGTTTGTCTGATTATTA 42506
Qy 73 ValAlaLeuAspValLysProValGlyLeuAlaAspAspAsnLysLysAsnArgIlelle 92
Db 42507 GCACATTAGATGAAACCTGTGTGTTATGACATGCGCAGCATGTAATAATAACA 42566
Qy 93 LysProLeuArgAspLysIleGlyLysThrSerValGlyThrArgLysGlnProAsn 112
Db 42567 AAGTCAGTAAGAGATAAGATTGGGGCATATGAATCGGTGGATCTAGACACCGAAT 42626
Qy 113 LeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsnAsnArgHisLys 132
Db 42627 ATGGAAGTGATAGTAATTAATAACCGGATTTGATCATTCGACATGTTAGCAGACATAAG 42686
Qy 133 GlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGly 152
Db 42687 AAAATCAATCAGAATTGAGCAAAATTCGCCGACCAATATTAGTCAGTGTGACGGGA 42746
Qy 153 AspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 172
Db 42747 GATTACAAATGCAATATTGGAAGCATTTAAACAGTCCTAGACATGTCGCAAGAGAAA 42806
Qy 173 GluGlyLysLysArgLeuGluGluHisAspLysLysIleGluLutyrLysLysGluIle 192
Db 42807 GAAGCGAGAGAGCGTCTGGAAAAGCATGATATAAATATTAGCGGAGATTAGAAAAGAAAT 42866
Qy 193 ThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAla 212
Db 42867 GAACAGAGTACGTTAAATATCGCAATTTGCAATTCGATATCTCAAGAGCAGGATGTTATT 42926
Qy 213 HisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232
Db 42927 AATAATGAAGATACATTATGGGCAATCTTACTTAAATGGGTATTCACACCTGAAATC 42986
Qy 233 SerAspAspValThrLysGlyLeuSerLysTyrLysLysGlyProTyrLeuGlnMetAsn 252
Db 42987 ACAAAGACAAAACCTACGCAATGTTGGTGAACGCAAGGGTGTCTTATATATATTTAAAT 43046
Qy 253 ThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn---LysAla 271
Db 43047 AATGAAGAACTTGCCTCAATATCAATCAAAAGTTATGATTTAGCCACTGACGGGAAACG 43106
Qy 272 SerSerAsnGluProSerLysLysGluLeuGluLysAspPro---ValTyrLysLysLeu 290
Db 43107 GACAAAATAAGACGAAATTC-----ATTGATCTCGCATGTTGGAAATCATTA 43154
Qy 291 AsnAlaValLysAsnGlnArgValAspIleLeuAspAspAspLeuTyrAlaArgSerArg 310
Db 43155 AAAGCTGTGAAAGATAACCAAGTTTATGCGTTGACGCAAAATAAGTGGTTGAAATCAAG 43214
Qy 311 GlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328
Db 43215 GGGATTATCGCAAGTGAAGATATGCGCAGAGATTTAGAAAATAATTCGAGAAAAA 43268

AP004825 290150 bp DNA linear BCT 20-DEC-2002
LOCUS Staphylococcus aureus subsp. aureus MW2 DNA, complete genome,
strain:MW2, section 4/10.
ACCESSION AP004825 BA000033
VERSION AP004825.1 GI:21203998
KEYWORDS
SOURCE Staphylococcus aureus subsp. aureus MW2
ORGANISM Staphylococcus aureus subsp. aureus MW2
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE
AUTHORS 1 Baba,T., Takeuchi,F., Kuroda,M., Yuzawa,H., Aoki,K., Oguchi,A.,
Nagai,Y., Iwama,N., Asano,K., Naimi,T., Kuroda,H., Cui,L.,
Yamamoto,K. and Hiramatsu,K.
Genome and virulence determinants of high virulence
community-acquired MRSA
Lancet 359 (9320), 1819-1827 (2002)
JOURNAL 22040717
MEDLINE 1204378
PUBMED
REFERENCE 2 (bases 1 to 290150)
AUTHORS Director-General, Biotechnology Center, Aoki,K., Oguchi,A.,
Nagai,Y., Asano,K., Iwama,N., Baba,T., Kuroda,M., Hiramatsu,K. and
Kikuchi,H.
Direct Submission
Submitted (06-MAR-2002) Director-General, Biotechnology Center,
National Institute of Technology and Evaluation, Biotechnology
Center; 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
(E-mail:bio.nite.go.jp, URL:http://www.bio.nite.go.jp/).
Tel:81-3-3481-1933, Fax:81-3-3481-8424)
Tel:81-3-3481-1933, Fax:81-3-3481-8424)
LOCATION/Qualifiers
1. 290150
/organism="Staphylococcus aureus subsp. aureus MW2"
/mol_type="genomic DNA"
/strain="MW2"
/db_xref="taxon:196620"
285..1493
/gene="MW0823"
285..1493
/gene="MW0823"
/note="ORFID:MW0823
hypothetical protein, similar to NADH dehydrogenase"
/codon_start=1
/transl_table=11
/protein_id="BAB94688.1"
/db_xref="GI:21203990"
/translation="MAQDRKKVLVLGAGYAGLVTKLQKAISTEAEITLKNVEV
YEATLHEAGAGTLNEDVLYPVESVLKKDKVQAEVTKIDRDKKVTETNQGIYDF
DILVVALGVSETIGIBGMKHAFIENVITARELSHEDIEKFAVAAASKEDKNDLSD
ILVAGGATGVETIGELTDPCLSKYGVQDKKITCVFAAPKMLPMFSELVNHA
VSYLEDGVEPKIATPIVACNEGFVVEVDEGKQNLNAGTSVMAAGVSGSKLMEESFE
GVKGRIVTKQDLTVGIDNI FVIGDCSAP I PAGEERPLPTTAQIAQQGSEVAKNIK
RILNGSETEFEYVDRTGTVCSLGHSDGVHVFDPKPIAGKKAAPMKKVIDTRAVFKIG
IGLAFKKGK"
1624..3099
/gene="ampA"
1624..3099
/gene="ampA"
/note="ORFID:MW0824"
/codon_start=1
/transl_table=11
/product="probable cytosol aminopeptidase"
/protein_id="BAB94689.1"
/db_xref="GI:21203991"
/translation="MNFKLNTLSNLTLLIGIPEHLNQLERISFNHIDITESLERL
KHQIHGKVGKIYTTAFDQDQYRLITVLGKLNKTPSYODMLKIWGHLFOYTKSEH
IDSTYLLMDSFISKYDQLSDVLMAAGIOSERATVEFPHYKSKKAPKTNLNLISL
IELDTFHGISIGISGINSARDFSNNPPNLPQTFAEDIVNHFKNTKKVDVQKDYDL
VSEGLGLQAGKSGKHPRLVITITNGKVDPEALVVGKGIITVSGSGYSIKTKNGST
ATMKVDMCAANVGIIEAASRLQIPVNIIVGLVACANINEASMKKPDQVFTALSGET
VEVMYTDAGRLVLADAFYANQYOPSPYMFATLTAAGIIVALGDKDAAPESNSKVI
LNDIQLI SSEVDVMVPELPITATERASIKHSDIADLVNHTNGQKALPAASFVTHFSG
QTPIHFDIAGPATTKNSYNGPKGTPTIVQWLKQQ"
3510..4826

CDS /gene="MW0825"
3510..4826
/gene="MW0825"
/note="ORFID:MW0825"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB94690.1"
/db_xref="GI:21203992"
/translation="MINAVYAVILMVLVLCRLNVVISIFISALVGGILSGMSIEKV
INVFGKNIVDGAVSALYALGGFPAALISVSGTIDVLGKIIINAIHAENSRMRVKV
VTIIALLAMSIQNLIPVHAFIPVIPPUSLNFNDKIDRRLLGLIIGFGLCPY
VLLPYFGQIQIIOGFAKANHPDEFNMWKAMLIIPSGYIVGLLIGLYVVRKPRE
YETRKISDSDNVELKPYLIVITVAIATFLVQTTDSMIFGALAGLVVFFTSRAY
WYELDAKFVEGIMAYIGVILTANGFAGVMNATGIDELVKLTLSITGDNKLPFII
MMVIGLITVLGIGSSFPATIPITASLIFPGFSGIGLDTWALIALIGTASALGSGSPA
SDSTLGTAGLVGGQHDHIRDTCVNFLEFYNIPLMIFGTIAMVL"
4845..5219
/gene="MW0826"
4845..5219
/gene="MW0826"
/note="ORFID:MW0826"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAB94691.1"
/db_xref="GI:21203993"
/translation="MTHLETPEMSIDHOEDGLVLSMPVTDKVKQPGCYLHGASIA
LGETAACLSGANLIDTTKEPIPLGLENNAIHSAKGDHVTATAEIHHQSKTHVWDIK
IKNDXEQLITVNRGTVAIKPLK"
complement(5275..6429)
/gene="MW0827"
complement(5275..6429)
/gene="MW0827"
/note="ORFID:MW0827"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB94692.1"
/db_xref="GI:21203994"
/translation="MHWIIGGIQGTIAQKLSSGTTDLRLTIIDPHETFCQRFNS
FTRNEMPLKSPVHHVHPQFHLQFAKHQYTNAPYQRPETLTFMFDHIAHAS
KOYLEDLQVLTLDKQEDKWHIKLEGGIITDCVIAIGSTNIPFMDILKDK
QADNFIPEKHDOVYDTHIVGSGITAAHLAKLLNHDNDKTHLWLNKDIEHDF
DADPCWGLPKMSTFLSKTSPERNAIQORERHKGSMPEHLYLRLKHGKGNRNVK
TPIQISGLINTENSVYQIIMVATGFEQDFMSQPLIKGLIQNYDAPINECNPVI
SEKLEWINPLVAGCFADLELGFGRNVNGRKAERIEQAFLKLOQYSA"
complement(6678..7034)
/gene="mnhg"
complement(6678..7034)
/gene="mnhg"
/note="ORFID:MW0828"
/codon_start=1
/transl_table=11
/product="Na+/H+ antiporter subunit"
/protein_id="BAB94693.1"
/db_xref="GI:21203995"
/translation="WIKILISLALIFVIGALISALAAAGLLBLEVDVYSRAHAGKA
STLGMSLLFETFLYFIATQGFVNNQLIAIFLITGFLSSHMIMKAAVNIKTPYTK
KTKVDEISEDLKDKTL"
complement(7012..7305)
/gene="mnhf"
complement(7012..7305)
/gene="mnhf"
/note="ORFID:MW0829"
/codon_start=1
/transl_table=11
/product="Na+/H+ antiporter subunit"
/protein_id="BAB94694.1"
/db_xref="GI:21203996"
/translation="MWNHVIIVITALLIIVISMLAMLRVIRVGLPSADRVVALDAIGLQ
LMVIALFSLILNKKIWMIVIMVIMIGILAFLTAVFSKFMKDGKVIHEDQNH"
complement(7305..7784)

CDS /gene="mnhE"
complement(7305..7784)
/gene="mnhE"
/note="ORFID:MW0830"
/codon_start=1
/transl_table=11
/product="Na+/H+ antiporter subunit"
/protein_id="BAB94695.1"
/db_xref="GI:21203997"
/translation="MAVQLVNFIIIAVFWLFTVNSVTNNFVLFGLVGLVLLHRV
LPGRFVYIYRIILKLVIIPLIELIKANFDVLKIIKPSIKNEPGFVYHTDLKDKHQ
IVLLSNLITLTCTVVLGVSDDRTKIYIHAIDPSTKEQVESIKTSLEKIVREVEGI"
complement(7786..9282)
/gene="mnhD"
complement(7786..9282)
/gene="mnhD"
/note="ORFID:MW0831"
/codon_start=1
/transl_table=11
/product="Na+/H+ antiporter subunit"
/protein_id="BAB94696.1"
/db_xref="GI:21203998"
/translation="MIESNMLVTLVPIVITALLVFIKRPILIKYVALGGTLLTIV
AAIINLVNKHGPIRVELGSKAPYSIVFDLIDFSAIIISIIITAILVSYOTI
GIERERYVYFVFLMGLIGIAGFTGIDFNLFVFEVFLMSSYFLVIVISYQIQOE
TIKYVLNVVSSFFVMGVAILYSVGTNLADISKNLANLSAHSGLVNIIVILFIF
VFATKAGFPMFVWLPISAYAPPPIIAFFGALITKVGYAIARTLSLFFSDNVSFH
VYILFALLITIFGCVGAVANIKKIILYNMIAVGVLGVAMTESGMICAIYVT
LHDMVLKALFLILGIMIKITGTLRQFGLIKRYPVLGWSFFTAALSLAGIPPLSG
FYGKFFIVQSTFERCFYLSGVILSSLVLSYIRIFLQGFQPGKGYDNNKVDVK
YLTTIAVAVITVLYGUSADLYLPMVKAGAEFFYNPSTYVKAIVLGGK"
complement(9275..9616)
/gene="mnhC"
complement(9275..9616)
/gene="mnhC"
/note="ORFID:MW0832"
/codon_start=1
/transl_table=11
/product="Na+/H+ antiporter subunit"
/protein_id="BAB94697.1"
/db_xref="GI:21203999"
/translation="WEIIMIFVSGILTAISVVLVSKSLIRIVMGTLLTHAANLFI
TMGKGVGPEIYIEANVKSVDPIQALILTAIVIAFATTAFFLAFRTYKELGTDN
VESMKGVPEDD"
complement(9616..10044)
/gene="mnhB"
complement(9616..10044)
/gene="mnhB"
/note="ORFID:MW0833"
/codon_start=1
/transl_table=11
/product="Na+/H+ antiporter subunit"
/protein_id="BAB94698.1"
/db_xref="GI:21204000"
/translation="MNRQNDLILQAAVIFFMVMVFGFSLFAGHYTPGGGFVGGI

Alignment Scores:

Pred. No.:	8.37e-43	Length:	290150
Score:	725.50	Matches:	149
Percent Similarity:	66.67%	Conservative:	63
Best Local Similarity:	46.86%	Mismatches:	95
Query Match:	42.93%	Indels:	11
DB:	1	Gaps:	4
US-10-724-972A-6352 (1-335) x AP004825 (1-290150)			
Qy	13	ValleGlyLeuLeuPheValLeuAlaThraAlaCysGlyAanAanSerSerSer	32
Db	108019	GTGTGTTTATGCTAATCTTAGTTGTAGCAGTCGGGTGTGCTCAAAAGACTACT---	108075
Qy	33	AanSerSerLysGluSerSerLysAapGlyValGluLeuLysHisGluGluGlyThrThr	52
Db	108076	-----GAAGAGAAACTGAATGACGACATAAAAGATGAATTAGGAACCTGAA	108123

JOURNAL		Patent: WO 02059148-A 301 01-AUG-2002;	
FEATURES		Cistem Biotechnologies GmbH (AT)	
Source		Location/Qualifiers	
		1..957	
		/organism="Staphylococcus aureus"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:1280"	
ORIGIN			
Alignment Scores:			
Pred. No.:		1.05e-45	Length: 957
Score:		724.50	Matches: 149
Percent Similarity:		66.67%	Conservative: 63
Best Local Similarity:		46.86%	Mismatches: 95
Query Match:		42.87%	Indels: 11
DB:		6	Gaps: 4
US-10-724-972A-6352 (1-335) x AX583924 (1-957)			
QY	13	ValileGlyLeuLeuPheValLeuileAlaThraAlaAalaCysGlyAenAenSerSerSer 32	
Db	25	GTGTGTTTCATGCTAATTTTGTAGTGTAGCGGTGTGTGTCCTCAAAAAGATACT- 81	
QY	33	AsnSerSerLysGluSerSerLysAspGlyValGluileLysHisGluGluGlyThr 52	
Db	82	-----GAAGAGAAAACTGAAATGACGACAAATAAAAGATGAATTAGGACTCAA 129	
QY	53	LysValProLysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeu 72	
Db	130	AAATTAAGAAAAATCCTAAACGTTGTTGTTGTTATTAGATATAGTTTGTGCTGATTATTA 189	
QY	73	ValAlaLeuAspValLysProValGlyileAlaAaspAspAenLysLysAenArgIleile 92	
Db	190	GCAGCATTAGATGAAACCTGTTGGTATTGCAGATGATGCGACACTAAATAATATAACA 249	
QY	93	LysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAen 112	
Db	250	AAATCAAAATCAGAAATGAGCAAAATGCTCCGCAATCATGTTAGTTCGGTACGGGA 429	
QY	113	LeuGluGluileSerLysLeuLysProAspLeuileileAlaAaspAenAenArgHisLys 132	
Db	310	ATGGAAGTGATAGTAATTAATTAACCGGATTTGATCATTGTCAGATGTTAGCAGACATAAG 369	
QY	133	GlyileTyrLysAspLeuAenLysileAlaProThrileGluLeuLysSerPheAspGly 152	
Db	370	AAATCAAAATCAGAAATGAGCAAAATGCTCCGCAATCATGTTAGTTCGGTACGGGA 429	
QY	153	AspTyrAenGluAenLysileAspAlaPheLysThrileSerLysAlaLeuGlyLysGluGlu 172	
Db	430	GATTATAATGCAAAATATTGAAGCATTTAAACAGCTGCTAAAGCAGTAGGCAAGAGAAA 489	
QY	173	GluGlyLysLysArgLeuGluHisAspLysLysileGluGluTyrLysLysGluile 192	
Db	490	GAAGCGGAGAGCGTCTGAAAAGCATGTTAAATATTAGCGGAGATTAGAAAAGAAATTT 549	
QY	193	ThrMetAspLysAenGlnLysValLeuProAlaValAlaAlaLysSerGlyLeuLeuAla 212	
Db	550	GAACAGAGTAGCTTAAATCTGCATTTGCAATTCGCTATCTCAAGAGCAGGTATGTTATT 609	
QY	213	HisProSerAenSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232	
Db	610	ATAATATGAAGATACATATTTATGGGACAAATCTTTAAATTAATGGGTATTCACCTGAAGTC 669	
QY	233	SerAspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAen 252	
Db	670	ACAAAGAGCAAAACTACGCATGTTGGTGAAACCGCAAGGTGGTCTTATATATATTAAT 729	
QY	253	ThrGluThrLysSerGlnValAenProGluArgMetPheIleMetThrAen---LysAla 271	
Db	730	AATGAAGAACTTGCCCAATATCAATCCAAAAAGTATGATTTAGCCACTACCGGAAAAACG 789	
QY	272	SerSerAenGluProSerLeuLysGluLeuLysAspPro---ValTyrLysLysLeu 290	

QY	53	LysValProLysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeu 72	
Db	108124	AAAATTAAAGAAAAATCCTAAACGTTGTTGTTATTAGAAATATAGTTTGTCTGATTATTTA 108183	
QY	73	ValAlaLeuAspValLysProValGlyileAlaAaspAspAenLysLysAenArgIleile 92	
Db	108184	GCAGCATTAGATGAAACCTGTTGGTATTTCAGATGATGCGACAGTAAATAATAATAACA 108243	
QY	93	LysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAen 112	
Db	108244	AAGTCAGTAAGAGATAGATTGGGGCATATGAATCGGTTGATCTAGACCGCAACCGAAT 108303	
QY	113	LeuGluGluileSerLysLeuLysProAspLeuileileAlaAaspAenAenArgHisLys 132	
Db	108304	ATGGAAGTGATAGTAATTAATTAACCGGATTTGATCATTCGAGATGTTAGACACATAAG 108363	
QY	133	GlyileTyrLysAspLeuAenLysileAlaProThrileGluLeuLysSerPheAspGly 152	
Db	108364	AAAATCAAAATCAGAAATGAGCAAAATGCGCGCAAAATATTGTTAGTCAGTGGTACGGGA 108423	
QY	153	AspTyrAenGluAenLysileAspAlaPheLysThrileSerLysAlaLeuGlyLysGluGlu 172	
Db	108424	GATTACAAATGCAATATTGAAGCATTTAAACACGTCGCTAAAGCAGTTGGCAAGAGAAA 108483	
QY	173	GluGlyLysLysArgLeuGluGluHisAspLysLysileGluGluTyrLysLysGluile 192	
Db	108484	GAAGCGGAGAGCGTCTGAAAAGCATGATAAAATATTAGCGGAGATTAGAAAAGAAATTT 108543	
QY	193	ThrMetAspLysAenGlnLysValLeuProAlaValAlaAlaLysSerGlyLeuLeuAla 212	
Db	108544	GAACAGAGTAGCTTAAATCTGCATTTGCAATTCGCTATCTCAAGAGCAGGTATGTTATT 108603	
QY	213	HisProSerAenSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232	
Db	108604	ATAATATGAAGATACATTTATGGGACAAATCTTACTTAAATGGGTATTCACCTGAAGTC 108663	
QY	233	SerAspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAen 252	
Db	108664	ACAAAGAGCAAAACTACGCATGTTGGTGAACGCAAGGGTGGTCTTATATATATTAAT 108723	
QY	253	ThrGluThrLysSerGlnValAenProGluArgMetPheIleMetThrAen---LysAla 271	
Db	108724	AATGAAGAACTTGCCCAATATCAATCCAAAGTTATGATTTAGCCACTACCGGAAAAACG 108783	
QY	272	SerSerAenGluProSerLeuLysGluLeuLysAspPro---ValTyrLysLysLeu 290	
Db	108784	GACAAAATAGAACGAAATTC-----ATTGATCCTGCAGTTTGGAAATCATTA 108831	
QY	291	AsnAlaValLysAenGlnArgValAlaAspIleLeuAspArgAspLeuTrpAlaArgSerArg 310	
Db	108832	AAAGCTGTGAAGATAACAAAGTTTATGAGTTGACCGAAATAAGTGGTTGAAATCAAGG 108891	
QY	311	GlyLeuileSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328	
Db	108892	GGGATTTCGCAAGTGAAGATGCGCAGAGATTTAGAAAAAATTCGCAAGAAA 108945	

RESULT 10	
AX583924	
LOCUS	957 bp DNA linear PAT 10-JAN-2003
DEFINITION	Sequence 301 from Patent WO2059148.
ACCESSION	AX583924
VERSION	AX583924.1 GI:27655594
KEYWORDS	Staphylococcus aureus
SOURCE	Staphylococcus aureus
ORGANISM	Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE	1
AUTHORS	Mainke, A., Nagy, E., von Ahsen, U., Klade, C., Henics, T., Zauner, W., Minn, D.B., Vytvytska, O., Eitz, H., Dryla, A., Weichhart, T., Hafner, M., Tempelmeier, B., Fraser, C.M. and Gill, S.
TITLE	A method for identification, isolation and production of antigens

QY	53	LysValProLysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeu 72	
Db	108124	AAAAATGAAGAAAAATCCTAAACGTTGTTGTTATTAGATATAGTTTGTGCTGATTATTA 108183	
QY	73	ValAlaLeuAspValLysProValGlyileAlaAaspAspAenLysLysAenArgIleile 92	
Db	108184	GCAGCATTAGATGAAACCTGTTGGTATTGCGATGATGCGACAGTAAATAATATAACA 108243	
QY	93	LysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAen 112	
Db	108244	AACTCAGTAAGAGATAGATTGGGGCATATGAATCGTTGGATGTAGACCGCAACCGAAT 108303	
QY	113	LeuGluGluileSerLysLeuLysProAspLeuileileAlaAaspAenAenArgHisLys 132	
Db	108304	ATGGAAGTGATGATAATTAATTAACCGGATTTGATCATTGCGATGTTAGCACACATAAG 108363	
QY	133	GlyileTyrLysAspLeuAenLysileAlaProThrileGluLeuLysSerPheAspGly 152	
Db	108364	AAAATCAAAATCAGAAATGAGCAAAATGCGCCGACAAATATTGTTAGTCACTGTTACGGGA 108423	
QY	153	AspTyrAenGluAenLysileAspAlaPheLysThrileSerLysAlaLeuGlyLysGluGlu 172	
Db	108424	GATTACANTGCAATATTGAAGCATTTAAACAGCTGCTAAAGCAGTTGCGCAAGAGAAA 108483	
QY	173	GluGlyLysLysArgLeuGluGluHisAspLysLysileGluGluTyrLysLysGluile 192	
Db	108484	GAAGCGGAGAGCGTCTGAAAAGCATGATAAAATATTAGCGGAGATTAGAAAAGAAATTT 108543	
QY	193	ThrMetAspLysAenGlnLysValLeuProAlaValAlaAlaLysSerGlyLeuLeuAla 212	
Db	108544	GAACAGAGTAGCTTAAATCTGCATTTGCAATTCGCTATCTCAAGAGCAGGTATGTTATT 108603	
QY	213	HisProSerAenSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232	
Db	108604	AATATGAAGATACATTTATGGGACAAATCTTACTTAAATGGGTATTAACCTGAAAGTC 108663	
QY	233	SerAspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAen 252	
Db	108664	ACAAAGAGCAAAACTACGCATGTTGGTGAACGCAAGGTTGCTCTTATATATATATTAAT 108723	
QY	253	ThrGluThrLeuSerGlnValAenProGluArgMetPheIleMetThrAen---LysAla 271	
Db	108724	AATGAAGAACTTGCCCAATATCAATCCAAAAAGTATGATTTAGCCACTACCGGAAAACG 108783	
QY	272	SerSerAenGluProSerLeuLysGluLeuLysAspPro---ValTyrLysLysLeu 290	
Db	108784	GACAAAAATAGAACCAATTC-----ATTGATCCTGCAGTTTGGAAATCATTA 108831	
QY	291	AenAlaValLysAenGlnArgValAspLeuLeuAspArgAspLeuTyrAlaAargSerArg 310	
Db	108832	AAAGCTGTGAAGATGAACAAAGTTTATGACGTTGACCGGAAATAGTGGTTGAATCAAGG 108891	
QY	311	GlyLeuileSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328	
Db	108892	GGGATTATCGCAAGTGAAAGTATGCGAAGAGATTAGAAAAAATTTGCAAAAAA 108945	

RESULT 10
AX583924
LOCUS
DEFINITION
Sequence 301 from Patent WO02059148.
ACCESSION
AX583924
VERSION
AX583924.1 GI:27655594
KEYWORDS
Staphylococcus aureus
SOURCE
Staphylococcus aureus
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE
1
AUTHORS
Meinke, A., Nagy, E., von Ahsen, U., Klade, C., Henics, T., Zauner, W.,
Minh, D. B., Vytvytska, O., Eitz, H., Dryla, A., Weichhart, T., Hafner, M.,
Tempelmeier, B., Fraser, C. M. and Gill, S.
TITLE
A method for identification, isolation and production of antigens
to a specific pathogen


```
Db 790 GACAAAATAGACAAATTC-----ATTGATCTCGCAGTTTGGAAATCATTA 837
Qy 291 AsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArg 310
Db 838 AAGCTGTGAAGATGAACAAAGTTTATGACGTTTGACCGAAATAGTGGTTGAAATCAAGG 897
Qy 311 GlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328
Db 898 GGGATTATCGCAAGTGAAGATGAGCAAGATTTAGAAAAAATTGCAGAAAAA 951

RESULT 11
AX621226
LOCUS 975 bp DNA linear PAT 20-FEB-2003
DEFINITION Sequence 4189 from Patent WO02094868.
ACCESSION AX621226
VERSION AX621226.1 GI:28449932
KEYWORDS
SOURCE Staphylococcus aureus
ORGANISM Staphylococcus aureus
REFERENCE 1
AUTHORS Masignani, V.C., Mora, M.C. and Scarselli, M.C.
TITLE Staphylococcus aureus proteins and nucleic acids
JOURNAL Patent: WO 02094868-A 4189 28-NOV-2002;
Chiron Spa (IT)
FEATURES
Location/Qualifiers
1..975
/organism="Staphylococcus aureus"
/mol_type="unassigned DNA"
/db_xref="taxon:1280"

ORIGIN
Alignment Scores:
Pred. No.: 1,07e-45 Length: 975
Score: 724.50 Matches: 149
Percent Similarity: 66.67% Conservative: 63
Best Local Similarity: 42.86% Mismatches: 95
Query Match: 42.87% Indels: 11
DB: 6 Gaps: 4

US-10-724-972A-6352 (1-335) x AX621226 (1-975)

Qy 13 ValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCysGlyAsnSerSerSer 32
Db 43 GTTGTTTCATGCTAATTTAGTTGAGCAGTAGCGGGTTGGTCAAAAAGATACT--- 99
Qy 33 AsnSerSerLysGlySerSerLysAspGlyValGluIleLysHisGluGluGlyThrThr 52
Db 100 -----GAAGAGAAACTGAATGACGACACATAAAGATGAATTAGGAACCTGAA 147
Qy 53 LysValProLysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeu 72
Db 148 AAAATTAAGAAAAATCCTAAACGTGTTGTTATTAGAAATATAGTTTTCGTGATTATTA 207
Qy 73 ValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLysLysAsnArgIleIle 92
Db 208 CGAGATTAGATATGAACCTGTTGGTATTGCAGATGTCGACACTAAATAATATAACA 267
Qy 93 LysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsn 112
Db 268 AAGTCAGTAAGACATAAGATTGGGCGCATATGAATCGTTGGTATCAGACCCACCGAAT 327
Qy 113 LeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsnAsnArgHisLys 132
Db 328 ATGGAAGTGAAGTAAGTAATTAATAACCGGATTGATCATTTGAGATGTTAGCAGACATAAG 387
Qy 133 GlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGly 152
Db 388 AAATCAATCAGAAATTGAGCAAAATTCCTCCGCAATCATGTTAGTTAGCGGTACGGGA 447
Qy 153 AspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 172
Db 448 GATTATAATCAATATTGAAGCATTTAAACACAGTCGCTAAGCAGTAGGCAAGAGAAA 507
```

```
Qy 173 GluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGluTyrLysLysGluIle 192
Db 508 GAAGCGCAGAGAGCGTCTGGAAGACGATGATAAATATTAGCGGAGATTAGAAGAAATTT 567
Qy 193 ThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAla 212
Db 568 GAACAGAGTAGCTTAAATCTGCATTTGCATTTCGGTATCTCAAGAGCAGGTATGTTATT 627
Qy 213 HisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232
Db 628 AATAATGAGATACATTTATGGGACAAATCTTAATTAATAATGGGTATTTCAACCTGAAGTC 687
Qy 233 SerAspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsn 252
Db 688 ACAAAAGACAAAACCTAGCGCATGTTGCTGAACGCAAGGTTGCTCTTATATATATTAAAT 747
Qy 253 ThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn---LysAla 271
Db 748 AATGAAGAACTTGGCAATATCAATCAAAAGTTATGATTTTAGCCACTGACGGAACG 807
Qy 272 SerSerAsnGluProSerLeuLysGluLeuGluLysAspPro---ValTrpLysLysLeu 290
Db 808 GACAAAATAGAACGAATTC-----ATTGATCTCGCAGTTTGGAAATCATTTA 855
Qy 291 AsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArg 310
Db 856 AAAGCTGTGAAGATGAACAAAGTTTATGACGTTGACCGAAATAAGTGGTTGAAATCAAGG 915
Qy 311 GlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328
Db 916 GGGATTATCGCAAGTGAAGTATGCGAGAAGATTTAGAAAAAATTGCAGAAAAA 969
```

```
RESULT 12
AP003132 298050 bp DNA linear BCT 11-JAN-2003
LOCUS Staphylococcus aureus subsp. aureus N315 genomic DNA, complete
DEFINITION genome, section 4/10.
ACCESSION AP003132 BA000018
VERSION AP003132.2 GI:14349174
KEYWORDS
```

```
SOURCE Staphylococcus aureus subsp. aureus N315
ORGANISM Staphylococcus aureus subsp. aureus N315
REFERENCE 1
AUTHORS Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I.,
Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Lian, J., Ito, T., Kanamori, M.,
Matsumaru, H., Maruyama, A., Murakami, H., Hoshino, A.,
Mizutani-Ui, Y., Takahashi, N.K., Sawano, T., Inoue, R., Kaito, C.,
Sekimizu, K., Hikakawa, H., Kuhara, S., Goto, S., Yabuzaki, J.,
Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C.,
Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H. and Hiramatsu, K.
Whole genome sequencing of methicillin-resistant Staphylococcus
aureus
Lancet 357 (9264), 1225-1240 (2001)
21311952
11418146
```

```
2 (bases 1 to 298050)
Director-General, Biotechnology Center, Aoki, K., Oguchi, A.,
Hosoyama, A., Nagai, Y., Kuroda, M., Hiramatsu, K. and Kikuchi, H.
Direct Submission
Submitted (30-JAN-2001) Director-General, Biotechnology Center,
National Institute of Technology and Evaluation, Biotechnology
Center: 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan
(E-mail: bioelite.go.jp, URL: http://www.bio.nite.go.jp/,
Tel: 81-3-3481-1933, Fax: 81-3-3481-8424)
On Jun 12, 2001 this sequence version replaced gi:13700734.
```

```
Location/Qualifiers
1..298050
/organism="Staphylococcus aureus subsp. aureus N315"
/mol_type="Genomic DNA"
/strain="N315"
/sub_species="aureus N315"
```

```
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
```

```
TITLE
JOURNAL
```

```
COMMENT
FEATURES
source
```



```

/db_xref="taxon:158879"
280_.432
/gene="SA0792"
CDS
280_.432
/gene="SA0792"
/notes="ORFID:SA0792"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB42031.1"
/db_xref="GI:13700735"
/translation="MKSLSQPNKYVEAFKPYLLTLLYLAIFITLYLYGSGDTHNN
FLYNEP"
448_.1905
/gene="dlcA"
CDS
448_.1905
/gene="dlcA"
/notes="ORFID:SA0793"
/codon_start=1
/transl_table=11
/product="D-alanine-D-alanyl carrier protein ligase"
/protein_id="BAB42032.1"
/db_xref="GI:13700736"
/translation="MTDILNQAFADANPQSIQAVRHTTDELTYQQLMDESSKLAHRL
QSGKPMTLFGHSPMYLIVMGIGAKGCGYVPVDTSPEDRIKMIINKVQPEFVNT
TDESFESEGVETIEDIKTSQDPVIFDSQIKONDVTVTFTSGTGEKGVQIEVAS
LVQFTHEMLELKNKGNGQQWLNAQPFSDLSVMAIYPCLASGGTLLNLVKNKMWKPKL
LNEMLTATPINIWSVTSFSEMCILLPTLNEEQVGSLSNEPFCCEILLPHRAAKLVSR
FSPATINYTYGPTATAVFTSDGIRTYHTGDAKPFENGQWFIQGRIDFQIKLNGYRME
QSVSLYLNKQKTAEVNFDGIRTYHTGDAKPFENGQWFIQGRIDFQIKLNGYRME
LEEITQLRQSEFVKEALVVPYKNDKVIHLIGAIVPTTEVTDNAENTKNIKNDLKR
LPEYIIPKFEWMEQLPLTNSGKIDRKKIAEVING"
1902_.3116
/gene="dlcB"
CDS
1902_.3116
/gene="dlcB"
/notes="ORFID:SA0794"
/codon_start=1
/transl_table=11
/product="DltB membrane protein"
/protein_id="BAB42033.1"
/db_xref="GI:13700737"
/translation="MIPYGDFTFFLIALLPVIILGLGKRSYIYNGVVTAFMVIL
IPSSDKHLFDQKYLISVOLISPIIVVMQVLLIMFYHSPKNNFSKFTVMVKSIL
PLALVKVQSTWLGQGHQHFHRESKLIEFVGLGISYVTFKSVQLIMEIRDGSIKVIV
WKLIQIFISFPTISQDLHGVDNWLMIYAYSILFFDPFAGYSLFALAFSYLFGIKTPPN
AYFINIYALPQDKFQFWRNWHMILSFWRDICIYMRSLFYMSRKKLLKSQFAMSNVAFILIN
FKKPPKAKNIKDFMWRNWHMILSFWRDICIYMRSLFYMSRKKLLKSQFAMSNVAFILIN
FTMGVTHGIEVYIIVGLYHAALFICYGYIERWRKKHPPRWQNGFTTALSIVITFHF
TGGLIFSGKLI"
3134_.3370
/gene="dlcC"
CDS
3134_.3370
/gene="dlcC"
/notes="ORFID:SA0795"
/codon_start=1
/transl_table=11
/product="D-alanyl carrier protein"
/protein_id="BAB42034.1"
/db_xref="GI:13700738"
/translation="MEFREQVLLLAEVAENDIVKENPDVEIFEGSIDSPOTVGLLI
EIONKLDIEVSIMDFORDEWATPNKIVEALEBLR"
3367_.4542
/gene="dlcD"
CDS
3367_.4542
/gene="dlcD"
/notes="(glycerophosphate chain)
ORFID:SA0796"
/codon_start=1
/transl_table=11
/product="poly D-alanine transfer protein"
/protein_id="BAB42035.1"
/db_xref="GI:13700739"

```

GAEEAPGENDEVLEYFGLKVLVDKEDAPVLNGTTIDPFQSLMGGGFQIDNPNAISCG
CGSSFRITAKVAGNPENC"
7973. 9181
/gene="SA0802"
7973. 9181
/gene="SA0802"
/notes="ORFID:SA0802
hypotheetical protein, similar to NADH dehydrogenase"
/codon start=1
/transl table=11
/protein_id="BAB42041.1"
/db_xref="GI:13700745"
/translation="MAQDRKKVLVAGYAGLQTVTKLQKAI STEBAEITL INKNEYH
YEATWLHEASAGLTNEDVLPVESVLKKDKNFVQAEVTKIDRDAKVETNGQIDYDF
DLVVALGFVSETFEGMKDHFQIENVITARELSHRHEDKFANYAAASKEDNDLS
ILVGGAGFTGVFELBTDRIPELCSKYGVQDNKVKITCEAPKPLMPFSELVNHA

Alignment Scores:
Pred. No.: 1,03e-42 Length: 298050
Score: 724.50 Matches: 149
Percent Similarity: 66.67% Conservative: 63
Best Local Similarity: 46.86% Mismatches: 95
Query Match: 42.87% Indels: 11
DB: 1 Gaps: 4

US-10-724-972A-6352 (1-335) x AP003132 (1-298050)

QY	13	ValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCysGlyAsnAsnSerSerSer 32
DB	113875	GTGTGTTTATGCTAATCTTAGTTGTAGCAGTAGCGGTTGTGTCACAAAAGATACT--- 113931
QY	33	AsnSerSerLysGlySerLysAspGlyValGluIleLysHisGluGluGlyThrThr 52
DB	113932	-----GAAGAGAAACTGAAATCAGCACAAATAAGATGAAATAGAACTGAA 113979
QY	53	LysValProLysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeu 72
DB	113980	AAATATTAGAAAAATCCTAAACGTGTGTGTATTAGATATAGTTTTCGCTGATTATTA 114039
QY	73	ValAlaLeuAspValLysProValGlyIleAlaAspAsnLysLysAsnArgIleIle 92
DB	114040	GCAGCATTAGATATGAACCTGTGTGTATGTGCAGATGATGCGCAGCACTAAAAATATAACA 114099
QY	93	LysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsn 112
DB	114100	AAGTCAGTAAGAGATAAGATTGGGGCATATGAATCGTTGGATCTAGACCGCAACCGAAT 114159
QY	113	LeuGluGluIleSerLysLeuLysProAspLeuIleIleAlaAspAsnAsnArgHisLys 132
DB	114160	ATGGAAGTCATGAATTAATTAACCGGATTTGATCATTTGCAGATGTTAGCAGACATAAG 114219
QY	133	GlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGly 152
DB	114220	AAATCAAAATCAGAAATGAGCAAAATTCCTCCGACAATCATGTTAGTTCGGGTACGGGA 114279
QY	153	AspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysLeuLeuGlyLysGluGlu 172
DB	114280	GATTATATGCAATATTGAAGCATTTAAACAGTCGCTAAAGCAGTAGGCAAGAGAAA 114339
QY	173	GluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGluTyrLysLysGluIle 192
DB	114340	GAAGCGGAGAGCGTCTGAAAGCATGATAAATATTAGCGGAGATTAGAAAGAAATTT 114399
QY	193	ThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAla 212
DB	114400	GAACAGAGTACGTTAAATAATCTGCATTTTCGATCTCAAGCAGCAGGTATGTTTATT 114459
QY	213	HisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232
DB	114460	AATTAATGAGATACATTTATGGCAATCTTAATTAATAATGGGTATTTCACACCTGCAAGTC 114519
QY	233	SerAspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsn 252

Db	114520	ACAAAAGACAAAACACTAGCCATGTTGGTGAACCAAGGTCCTCTATATATATTTAAAT 114579
QY	253	ThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn---LysAla 271
Db	114580	AATGAAGAACTTCCCAATATCAATCCAAAGTTATGATTTTAGCCACTGACGGAACG 114639
QY	272	SerSerAsnGluProSerLeuLysGluLeuGluLysAspPro---ValTrrLysLysLeu 290
Db	114640	GACAAAATAAGAACGAAATTC-----ATTGATCCTGCAGTTTGCAATCATTA 114687
QY	291	AsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrrAlaArgSerArg 310
Db	114688	AAAGCTGTGAAGATACAAAGTTTATGAGCTTGACCGGAATAAGTGGTTGAATCAAGG 114747
QY	311	GlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328
Db	114748	GGGATTATCGCAAGTGAAAGTATGCGAAGATTAGAAAAAATTCGACAGAAA 114801

RESULT 13
AP003361
LOCUS
DEFINITION
Staphylococcus aureus subsp. aureus Mu50 DNA, complete genome,
section 4/9.
ACCESSION
AP003361 BA000017
VERSION
AP003361.3 GI:46395333
KEYWORDS
Staphylococcus aureus subsp. aureus Mu50
SOURCE
Staphylococcus aureus subsp. aureus Mu50
ORGANISM
Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE
1

AUTHORS
Kuroda,M., Ohta,T., Uchiyama,I., Baba,T., Yuzawa,H., Kobayashi,I.,
Cui,L., Oguchi,A., Aoki,K., Nagai,Y., Lian,J., Ito,T., Kanamori,M.,
Matsumaru,H., Maruyama,A., Murakami,H., Hosoyama,A.,
Mizutani-Ui,Y., Takahashi,N.K., Sawano,T., Inoue,R., Kaito,C.,
Sekimizu,K., Hirakawa,H., Kuhara,S., Goto,S., Yabuzaki,J.,
Kanehisa,M., Yamashita,A., Oshima,K., Furuya,K., Yoshino,C.,
Shiba,T., Hattori,M., Ogasawara,N., Hayashi,H. and Hiramatsu,K.
Whole genome sequencing of methicillin-resistant Staphylococcus
aureus
Lancet 357 (9264), 1225-1240 (2001)

JOURNAL
MEDLINE
21311952
PUBMED
11418146
REFERENCE
2
Ohta,T., Hirakawa,H., Morikawa,K., Maruyama,A., Inose,Y.,
Yamashita,A., Oshima,K., Kuroda,M., Hattori,M., Hiramatsu,K.,
Kuhara,S. and Hayashi,H.
Nucleotide Substitutions in Staphylococcus aureus Strains, Mu50,
Mu3, and N315
DNA RESEARCH 11, 51-56 (2004)

JOURNAL
REFERENCE
3 (Bases 1 to 347785)
Direct Submission
Submitted (28-FEB-2001) Masahira Hattori, Kitasato Institute for
Life Sciences, Kitasato University; Kitasato 1-15-1, Sagamiara,
Kanagawa 228-8555, Japan
(E-mail:hattori@genome.its.kitasato-u.ac.jp, Tel:81-42-778-8194,
Fax:81-42-778-8193)
On Apr 15, 2004 this sequence version replaced gi:14246761.
This work was done in collaboration with Toshiko Ohta, Mutsumi
Kanamori, Hideo Hayashi (University of Tsukuba), Hideki Hirakawa,
Satoru Kuhara (Kyushu University), Ikuro Uchiyama (Okazaki National
Research Institutes), Susumu Goto, Minoru Kanehisa (Kyoto
University), Naotake Ogasawara (Nara Institute of Science and
Technology), and supported by the Research for the Future Program of
the Japan Society for the Promotion of Science.

LOCATION/Qualifiers
1. 347785
/organism="Staphylococcus aureus subsp. aureus Mu50"
/mol_type="genomic DNA"
/strain="Mu50"
/sub_species="aureus Mu50"
/db_xref="taxon:158878"
11. .997

FEATURES
source
1. 347785
/organism="Staphylococcus aureus subsp. aureus Mu50"
/mol_type="genomic DNA"
/strain="Mu50"
/sub_species="aureus Mu50"
/db_xref="taxon:158878"
11. .997

gene

gene
CDS
/codon_start=1
/locus_tag="SAV0992"
11..997
/locus_tag="SAV0992"
/codon_start=1
/transl_table=11
/product="oligopeptide ABC transporter ATP-binding protein homologue"
/protein_id="BAB57154.1"
/db_xref="GI:14246762"
/translation="MNNVLELVKDLSTLKNINWLAATVENISFELSKGEVLGIVGES GCGKSLISLIIKLLPEKISLSSGEVIFDGRKIDTLEKQLDIDRGNDAIMIFQSPM TALNPVTIKNOIVESI KSHKISKEANNLAKDLKKVGIARQDEILANSYHQLSGG MRQVIMAMASCPKLLIADEPTTALDVTIOAIDLLKELOKETQMAIMMITHDLS VVAEFCDKLVYAGQIVFEGGKEILHNPKHYPYTKLLSTIPKLEGGOKRLTIBGI VPSIQAFHNKCFANRCKKLDICNNQSPMHVCEVDIVRCHLYKNEYKEI"
1000..1980
/genes="appP"
/locus_tag="SAV0993"
1000..1980
/genes="appP"
/locus_tag="SAV0993"
/codon_start=1
/transl_table=11
/product="oligopeptide transport system ATP-binding protein AppP homologue"
/protein_id="BAB57155.1"
/db_xref="GI:14246763"
/translation="MENILEVNOIKKYKI KTGLLQTKYVKAVDVVSFSIKKGQTFG LVGSGCGKSTLGVIRLEDATSGSIIIVNGEDITRLQGGKKRSQYQYOMIFQDPYA SLNPMQVGIISPIILNYKKLPKEIKKEVLYLLKCVGLSEDAYYKYAEHFGSGQQR RVGARALALRPSIIVADEPVSALDVSVQSVOLNLLQLOEQKLSYLFIAHDLSVVK HISDVIQVYMLGHVIEASDKIYEENPAPHYTKALISSIPOIDKHNRRILKGLPSS PSNPQCPFHTRCPIAKDKCKENIPOLKDIGNEHQVACFYVDKVGDLNA"
1973..2935
/genes="oppB"
/locus_tag="SAV0994"
1973..2935
/genes="oppB"
/locus_tag="SAV0994"
/codon_start=1
/transl_table=11
/product="probable oligopeptide transport system permease protein"
/protein_id="BAB57156.1"
/db_xref="GI:14246764"
/translation="MLKILKRLGLMIPLLIISIVFSLAIQDPFSLQNGKIK QEANAQKEGLNDISI PHOYRWVNHVHIGDLGESIKYKRPVIDVIERIPNTILLG AMLSLIITYISPALGITSGRYSYSLDVTVOIFNYLMLAIPSIAGVFAIFISFELQ WPPQGSVDINIKEGTFPYNSKIYHTELPALTGLLSTAGVIOYLRNDIENSCKDY VLTARSKGLSMNKIYNKHLRNLSPITITFFLGADIVSILGAVITETIFSNGICKLF LESVTCQDYPLMMALTLPFSFLGLGNLISDITYGFIIDPRNSN"
2947..3828
/locus_tag="SAV0995"
2947..3828
/locus_tag="SAV0995"
/codon_start=1
/transl_table=11
/product="similar to oligopeptide ABC transporter permease"
/protein_id="BAB57157.1"
/db_xref="GI:14246765"
/translation="MQNKSQSPFKTALSFIHNKIAMLSIIFLLIITVISIAPLIAP FPNQDNLNIGEMTAQNLIGTSGGRDNFRLLYAGRISLSIGTITGMLLIGIT VGIISGYFGGI VDTLLMRTFEMFLPFLIFAIVLNAAIGDKIKNPYGSAILVLVII VLSWGIISGFGIIVKLVQLEKENEVFLAAKSIGTPTYKIILKLLPNLSVIVQATLLF AGHIVVESGLSFLGFGISKAI P8WGNMLSDAQEGDVI SGKPMWMPAIMITLILSI NFVGEGLKDAFNPRGRR"
3870..4859
/genes="trpS"
/locus_tag="SAV0996"
complement(3870..4859)
/genes="trpS"
/locus_tag="SAV0996"
complement(3870..4859)
/locus_tag="SAV0996"

gene
CDS
/codon_start=1
/transl_table=11
/product="tryptophanyl-tRNA synthetase"
/protein_id="BAB57158.1"
/db_xref="GI:14246766"
/translation="METLPSGIQPSGIPTIGNVIGALKOFVDVNDYDCYFICVDQHA ITMPQRLKRLQRLQAAIYLASGIDPDKATILFQSEVPAHVQAGMMITTIASVCEL ERMTOYKDKAQAQVEGIPAGLLTYPLMAADIVLNTNINVPVGGDDQKHIELTRNLVD RNSRYNDVLVKAQVEIPKVRMSLQDPTRMKSDNNAKNFISLLDEPNVAAKRI KAFATSDGIGIIPKDRNKPQITNLISYAGLTDMPIKOTI EAKYEGEGYKGFQGLAEI VKAFLVFEQEKYESFYNSDKLDDIILDQGRDKAHKVSFKTVKKMEKAMGLGRK"
5153..5548
/locus_tag="SAV0997"
5153..5548
/locus_tag="SAV0997"
/codon_start=1
/transl_table=11
/product="arsenate reductase family protein"
/protein_id="BAB57159.1"
/db_xref="GI:14246767"
/translation="MVTLTSPSCTSCRKAKAWLQEHDIPTYTERTNIFSEHLTIDEIKQ ILKMTEDGTDEIISTRSKYQKLNVDISLPLQDLYSIIQDNFGLLRPIIDNKRILQ VGNDEIRARFLPRKYRTFQLQEAQRMD"
5919..6638
/locus_tag="SAV0998"
5919..6638
/locus_tag="SAV0998"
/codon_start=1
/transl_table=11
/product="similar to negative regulator of genetic competence Meca"
/protein_id="BAB57160.1"
/db_xref="GI:14246768"
/translation="MRIERVDVTVTKLFIYTSIDTEARGFSGREDLWNRKGEPEFWSM MDINEBEDFVVEGPIWOVHAPEKGEVETISKSNEDMMNSDDATDQFQEQVQL LAQTEGEDQLEELFEORTKEKAQSKRQKSSARKNTRTIIVKFNLEDEVINYAHS NPITFEFDLLYMVDTGYIYVAFDHSVQDEVINDSYSLLEFPAYPTDRTEVYLNDYA KIMSHNVTAQVRRYPPEY"
6759..7745
/locus_tag="SAV0999"
6759..7745
/locus_tag="SAV0999"
/codon_start=1
/transl_table=11
/product="similar to transcription factor"
/protein_id="BAB57161.1"
/db_xref="GI:14246769"
/translation="MLVALNEEKERVLTATLRKTOYFCVCGKQVILKRLGVLSHF AHKLABQKCFNNETIKHYKSLILAOIQOQCGKVEIEPLKRIKOIPDILINNYY IELQYSPITPYKQIILQRTGELKMGYKSWLNDVDYCHNKVFNHFSLSLFINPIRKL HTFNLEKQIMMFQIQYLGHHKYVAEKRNAKIIELFNEAPCDHAYVYKSKFPAINQY IKYCRWNSVLEPTLSAMYQLTDQEVVYNYGIFFQIYIENHPIEWQLQVDLMLK NGSKLVNDNLNYFKLKKFIVALESKTAIEKLINNYLNICSDRGNVDQIILF"
7793..9601
/locus_tag="SAV1000"
7793..9601
/locus_tag="SAV1000"
/codon_start=1
/transl_table=11
/product="chimeric oligopeptidase homologue"
/protein_id="BAB57162.1"
/db_xref="GI:14246770"
/translation="MSQQLSRBEERKPYETWDLTTIPKDDAEFAAPKFEVENEIGK BEQFGHIGDSAEITLYNALLELEDTLTGLEKVVYVAHLKQDDTNDKTYGMSRAHQ LIKFSWASFLVPEILDEIDKIOSFVNSYDKLQKAFDLKLINEKPHILDAETEK LITEAQDALSTPSNYGMSNADLVFEDAIKQGNAPHLTQCTFIKYLSDSDRKLRES AFNRYKAYGAHNNTLGATLAGEVKNVFNATHNYKTAREKALSNHII PENYONLV KTVHKYPLLLHYTELRKELLGLDKMVDYLTPLIKDIFKEMPEEAEKEMWLKALEP MGEEYLVNVKGLNNRWVDYENKGRSGYSGAHLNPFIFLLNWSNTISDLYTLVH EFGHSASVFSKFPQSSNDYTI FVAEVAETCNBALLSDYMDKHLDDKRLLLNQE LERFRATLPRQTHAEFEHKAIEAGELPTPTMNEEYAKLNKLYFGDSVETDEDI SKWRSLRPHFYNNYIYQVATGYSAAQSLSHOILTEGKPAVDRIYINEFLKGGSSNYP I EILKNAGVDMTTPTEPEIQACEVFEQKLNAFEKLMKA"

gene

complement (10061. .10846)

Alignment Scores:
Pred. No.: 1.24e-42 Length: 347785
Score: 724.50 Matches: 149
Percent Similarity: 66.67% Conservative: 63
Best Local Similarity: 46.86% Mismatches: 95
Query Match: 42.87% Indels: 11
DB: 1 Gaps: 4

US-10-724-972A-6352 (1-335) x AP003361 (1-347785)

QY 13 ValileGlyLeuPheValLeuAlaThrAlaCysGlyAsnAsnSerSer 32

Db 46237 GTTCGTTTATGCTAATCTTGTAGTACGAGTGGGTTGTGCTCAAAAGATACT--- 46293

QY 33 AsnSerSerLysGlySerSerLysAspGlyValGluIleLysHisGluGluGlyThrThr 52

Db 46294 -----GAAGAGAAACTGAATGACGACCAATAAAGATGATGATAGAACTGAA 46341

QY 53 LysValProLysHisProLysArgValValValValLeuGluTyrSerPheValAspAlaLeu 72

Db 46342 AAAATTAGAAAAATCCTAAACGTTGTTGTATTAGAAATATAGTTTGTCTGATTATTA 46401

QY 73 ValAlaLeuAspValLysProValGlyLeuAlaAspAsnLysLysAsnArgIleIle 92

Db 46402 GCAGCAATTAGATGAAACTGTGTGTTATGTCAGATGTCGACACTAAAAATATAACA 46461

QY 93 LysProLeuArgAspLysIleGlyLysThrThrSerValGlyThrArgLysGlnProAsn 112

Db 46462 AAGTCAGTAAGAGATAAGATTGGGGCATATGATCGTTGGATCTAGACCGCAACCGAAT 46521

QY 113 LeuGluGluLysSerLysLysLysProAspLeuIleAlaAspAsnAsnArgHisLys 132

Db 46522 ATGGAAGTGAATGAATTAATAACCGGATTTGATCATTCAGATGTTAGCAGACATAG 46581

QY 133 GlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLysSerPheAspGly 152

Db 46582 AAAATCAATCAGAAATGAGCAAAATGCTCCGACAATCATGTTAGTCGCGTACGGCA 46641

QY 153 AspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 172

Db 46642 GATTATAATGCAATATTGACATTTAAACAGTCGCTAAAGCAGTAGGCAAGAGAAA 46701

QY 173 GluGlyLysLysArgLeuGluGluHisAspLysIleGluGluTyrLysLysGluIle 192

Db 46702 GAAGCGAGAGCGCTCTGAAAGCATGATATAAATATTAGCGGATATTAGAAAGAAAT 46761

QY 193 ThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAla 212

Db 46762 GAACAGAGTACGTTAAATCTGCATTTTCATTCGGTATCTCAAGAGCAGGTATGTTATT 46821

QY 213 HisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232

Db 46822 AATAATGAGATACATTTATGGGCAATCTTAAATTAATGGGTATTCAACCTGAAGTC 46881

QY 233 SerAspAspValThrLysGlyLysSerLysTyrLysLysGlyProTyrLeuGlnMetAsn 252

Db 46882 ACAAAAGACAAAACCTACGCATGTTGGTGAACCGCAGCGGTGCTCTATATATTAAT 46941

QY 253 ThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn---LysAla 271

Db 46942 AATGAAGAATCTTGCCAATATCAATCCAAAGTTATGATTTTACGCTACGCGAAACG 47001

QY 272 SerSerAsnGluProSerLeuLysGluLeuGluLysAspPro---ValTrpLysLysLeu 290

Db 47002 GACAAAATAGACGAAATTC-----ATTGATCTCGAGTTTGGAAATCATTA 47049

QY 291 AsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArg 310

Db 47050 AAAGCTGTGAAGATACAAAGTTTATGACGTTGACCAAAATAGTGGTGAATCAAGG 47109

QY 311 GlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db

Db	47110	GGGATTATCGCAAGTGAAGATATGGCAGAAAGATTAGAAAAAATTGCAGAAAAA 47163	
<p>RESULT 14 A64853/c LOCUS DEFINITION Sequence 10 from Patent WO9731114. ACCESSION A64853 VERSION A64853.1 GI:4530844 KEYWORDS SOURCE unidentified ORGANISM unclassified. REFERENCE 1 AUTHORS Burnham, M.K. and Hodgson, J.E. TITLE POLYNUCLEOTIDES AND AMINOACID SEQUENCES FROM STAPHYLOCOCCUS AUREUS JOURNAL Patent: WO 9731114-A 10 28-AUG-1997; SMITHKLINE BEECHAM PLC (GB) FEATURES Location/Qualifiers 1..2247 /organism="unidentified" /mol_type="unassigned DNA" /db_xref="taxon:32644"</p>			
ORIGIN			
<p>Alignment Scores: Pred. No.: 5.35e-45 Length: 2247 Score: 721.00 Matches: 169 Percent Similarity: 73.48% Conservative: 36 Best Local Similarity: 60.57% Mismatches: 69 Query Match: 42.66% Indels: 9 DB: 6 Gaps: 1</p>			
US-10-724-972A-6352 (1-335) x A64853 (1-2247)			
QY	1	GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeu 20	
Db	831	CGAGTGGGACNGATGAGAGGCTCTAAAAACTTTTGTATATTGGGATTATATGTCCTTA 772	
QY	21	IleAlaThrAlaAlaCysGlyAsnAsnSerSerSerAsn-SerSerLysGluSerSerly 40	
Db	771	CTTTTGTGTCAGCTTGTGTTAATACGGTAAATCAAGTAAAGAAAGATCATCAACATA 712	
QY	40	sAspGlyVal-GluIleLysHisGluGluGlyThrThrLysValProLysHisProLysA 60	
Db	711	AGATACATATTTCGGGTAAAGATGAAATGATGATGATGATGATGATGATGATGATGAT 652	
QY	60	rgValValLeuGluTyrSerPheValAspAla---LeuValAlaLeuAspValLysP 79	
Db	651	GTATCGTTGTATTAGAGTACTCATTTGCAAGGTTGCATTATCAGCATATATGCCGTTANAC 592	
QY	79	ro-ValGlyIleAlaAspAspAsnLysLysAsnArgIleIleLysProLeuArgAspLys 98	
Db	591	CAAGTTCGTATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 533	
QY	99	IleGlyLysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLys 118	
Db	532	TTGGGGGT-AATACTCTGTAGGTACACGTAA-CAAGCNAACCTTAAGAGANATTAGTAAA 475	
QY	119	LeuLysProAspLeuIleIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeu 138	
Db	474	TTAAANCCGGGATTAATATATCTCNCCTAGTAGACATAGAGTAAAGGTTATTAATAAGAATTA 415	
QY	139	AsnLysIleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIle 158	
Db	414	GCCCACTTGCACACA-TTATCATTTAAAGAGTTTGTATGGAGACTACACCCCAATATT 356	
QY	159	AspAlaPheLys-ThrIleSerLysAlaLeuGlyLysGluGluGluGlyLysArgLe 178	
Db	355	AATTCGTTCACCAACAATTCGTTAAAGCTTTAAATAAGANAAGAGAGCGGNAACGCTCT 296	
QY	178	uGluGluHisAspLysLysIleGluTyrLysLysGluIleThrMetAspLysAsnG1 198	

295 TGCTGAGCATGATAAAATTAATCAATAGTATAGAGATGAAATTAATTTAGTAGAAATCA 236
Db
198 nlysvValLeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTy 218
Qy
235 CAAAGTGCCTCCAGCAGTAGTCTTAAGAGCTGCTTTATAGCACATCCACACTATTCA 176
Db
218 rValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrly 238
Qy
175 TGTGGACAAATTTTAAACGACNCTAGGATTTAAATAATGCATTAAGTGACGATGAACACA 116
Db
238 sGlyLeuSerLysTyLeuLysGlyProTyLeuGlnMetAsnThrGluThrLeuSerGl 258
Qy
115 AGGTTAAAGTAAATATTGAAGAGCCTTACTTACAAATTAGACACACTGAACATTAGCTGA 56
Db
258 nValAsnProGluArgMetPheLeuMetThrAsnLysAlaSerSerAsn 274
Qy
55 TTTAAATCCAGAGCGTATGATCATTTAGACAGATCATGCTTAAATAAGAT 7
Db

RESULT 15
BX571856_10
WPCOMMENT
Sequence split into 29 fragments LOCUS BX571856 Accession BX571856

Fragment Name	Begin	End
BX571856_00	1	110000
BX571856_01	100001	210000
BX571856_02	200001	310000
BX571856_03	300001	410000
BX571856_04	400001	510000
BX571856_05	500001	610000
BX571856_06	600001	710000
BX571856_07	700001	810000
BX571856_08	800001	910000
BX571856_09	900001	1010000
BX571856_10	1000001	1100000
BX571856_11	1100001	1210000
BX571856_12	1200001	1310000
BX571856_13	1300001	1410000
BX571856_14	1400001	1510000
BX571856_15	1500001	1610000
BX571856_16	1600001	1710000
BX571856_17	1700001	1810000
BX571856_18	1800001	1910000
BX571856_19	1900001	2010000
BX571856_20	2000001	2110000
BX571856_21	2100001	2210000
BX571856_22	2200001	2310000
BX571856_23	2300001	2410000
BX571856_24	2400001	2510000
BX571856_25	2500001	2610000
BX571856_26	2600001	2710000
BX571856_27	2700001	2810000
BX571856_28	2800001	2902619

Continuation (11 of 29) of BX571856 from base 1000001 (BX571856 Staphylococcus aureus su

Alignment Scores:
Pred. No.: 7.38e-43 Length: 110000
Score: 719.50 Matches: 147
Percent Similarity: 66.67% Conservative: 65
Best Local Similarity: 46.23% Mismatches: 95
Query Match: 42.57% Indels: 11
DB: 4 Gaps: 4

US-10-724-972A-6352 (1-335) x BX571856_10 (1-1100000)
Qy 13 ValileGlyLeuPheValLeuPheValLeuAlaThrAlaAlaCysGlyAsnAsnSerSer 32
Db 54238 GTTGTGTTTATGCTATCTAGTTGTAGCAGTAGCGGTTGTGCTCAAAAGATACACT--- 54294
Qy 33 AsnSerSerLysGlySerSerLysAspGlyValGluLeuLysHisGluGluGlyThrThr 52
Db 54295 -----GAAGAGAAACTGAAATGACGACGATGAAGATGAATGAGTAACTGAA 54342
Qy 53 LysValProLysHisProLysArgValValValLeuGluThrSerPheValAspAlaLeu 72

54343 AAAATTAAGAAAAATCCGAAACGTATTGTTGATTAGAAATATAGTTTCTGCTATTATTTA 54402
Db
73 ValAlaLeuAspValLysProValGlyLeuAlaAspAspAsnLysLysAsnArgIlelle 92
Qy
54403 GCAGCATTAGATGAACCTGTTGGTATTGCGATGATGGTAGTACTAAAAATATAACA 54462
Db
93 LysProLeuArgAspLysIleGlyLysTyThrSerValGlyThrArgLysGlnProAsn 112
Qy
54463 AAGTCAGTAAAGATAAGATTGGGGCATATGAATCGTTGGATCTAGACCGCAACCGAAT 54522
Db
113 LeuGluGluIleSerLysLysLeuLysProAspLeuIleAlaAspAsnAsnArgHileLys 132
Qy
54523 ATGGAAGTGTATAGTAATTAATAACCGGATTGATTATTGCAGATGTCAGCAGACATAAG 54582
Db
133 GlyIleTyLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGly 152
Qy
54583 AAAATCAAATCAGAAATTGACAAAATTGGCCGACCAATCATGTTAGTCGGTACCGGA 54642
Db
153 AspTyAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 172
Qy
54643 GATTACAATGCAAAATATTGAAGCATTTAAACACAGTCTCTAAACAGATTGGCAAGAGAAA 54702
Db
173 GluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGluTyLysLysGluIle 192
Qy
54703 GAAGCGCAAAAGCGTCTGGAAGAGCATATAAATAATTAGCGGAGATTAGAAAAGAAATT 54762
Db
193 ThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAla 212
Qy
54763 GAACAGAGTAGCTTAAATCTGCATTTGCTATCTCAAGACGAGGTATGTTTATT 54822
Db
213 HisProSerAsnSerTyValGlyGlnPheLysSerGlnLeuGlyPheLysGluAlaLeu 232
Qy
54823 AATAATGAAGATACATTTATGGGACAATTCTTAATTAATAATGGGTATTCAACCTGAAGTC 54882
Db
233 SerAspAspValThrLysGlyLeuSerLysTyLysLysGlyProTyLeuGlnMetAsn 252
Qy
54883 ACAAAGACAAAACCTGGCATCTGTTGGTGAACGCGAGGGTGGCCCTTATATTATTATAAT 54942
Db
253 ThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn---LysAla 271
Qy
54943 AATGAGAGCTTGCCAATATCAATCCAAAAGTTATGATTTTAGCTAGCAATGGAACAAACA 55002
Db
272 SerSerAsnGluProSerLeuLysGluLeuGlyLysAspPro---ValTyLysLysLeu 290
Qy
55003 GATAAAAATAGAACGAAATTC-----ATTGATCTCGCAGTTTGGAAATCATTA 55050
Db
291 AsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArg 310
Qy
55051 AAAGCTGTGAAGATAATAAAGTATATGATGTCGAAATAGTGGTTGCAATCAAGA 55110
Db
311 GlyLeuIleSerSerGluMetAlaLysGluLeuValGluLeuSerLysLys 328
Qy
55111 GGTATTATGGCAAGTGAAGTATGCGAGAAGATTTAGAAAAAATTCAGAAAAA 55164
Db

RESULT 16
AX951891/c
LOCUS AX951891 5718 bp DNA linear PAT 08-JAN-2004
DEFINITION Sequence 20 from Patent WO03095658.
ACCESSION AX951891
VERSION AX951891.1 GI:40782280
KEYWORDS
SOURCE Bacillus subtilis
ORGANISM Bacillus subtilis
REFERENCE 1
AUTHORS Bj Rnvad, M.E., J Rgensen, P.L. and Hansen, P.K.
TITLE Homologous recombination into bacterium for the generation of polynucleotide libraries
JOURNAL Patent: WO 03095658-A 20 20-NOV-2003;
FEATURES
Novozymes A/S (DK)
Location/Qualifiers
1. .5718
source

[illegible]


```

SVASAREPIPPAITEALIMEISFEALREAGVRLPKTIQGTQVSIILGALVIGTAANVEAG
IVSAPMVIIVSITIGTASFTIIPRFNLAISIRMLRFLPMLFLASIFGIFGIMLGTIIVLVH
LCKLQSGFIPILSGISPFKXDEVKOLFVRAPWWTTRPGTYSRGNQKGAKREDPKD
BENNI"
RBS      3166.. .3185
CDS      3183.. .4337
/codon_start=1
/transl_table=11
/product="yfkR"
/protein_id="BAA22291.1"
/db_xref="GI:2443220"
/translation="MKKTIYKCVLPILLICLLTGCWDRTEINDIAFVVSASIDKKKKQ
YVAMQIPLVGGGTGGGGTAGSKTWVDSASGTTIREANNKLTQSLRSTINTSH
RRTVILGEMARDGAPVFDILTRNQBLTALILVSRGEARDILNTDVQLEOFPAEM
IRELAIITSDFVFLRMSDLVIGSDYPAPVISASKYKPGKGSKNLKIDGLAIFK
KRLMDIFKDEHMTAALMLNQARQPEFIVDLPNQMGQASIQLOKSNASFHAERKNGK
LSMTIRAKIIMENQSTVETRENDQYIIQKALNRTIKKOVISTVHRLQKUKADPA
GFODRTIRSTATTNLLKKEWEVYKMEVHVVPVITIEQGVLYKTISH"
RBS      4336.. .4355
CDS      4353.. .4553
/codon_start=1
/transl_table=11
/product="yfkS"
/protein_id="BAA22292.1"
/db_xref="GI:2443221"
/translation="MISYIVQTLIVCIAIYAEWKNFRSANNLTKNWAFSLIAGSAFL
WIYRVNPLPRLGHLPKYIPF"
RBS      4563.. .4579
CDS      4577.. .5653
/codon_start=1
/transl_table=11
/product="yfkT"
/protein_id="BAA22293.1"
/db_xref="GI:2443222"
/translation="MDKTSAYQGLFGALYTLAVGLKHAPILMTESAKQNAHWSYILG
WVIVPALMHRMLMKHQDKNIYELLSDSPSIAGRIIILFSLYFLILNAHDIRFI
NLINILPFPMAVLGVIIFVAICAREGKETLTPMAQIFLPFGILVLFPLFTLA
TQIELQNTIPFVEGLIIPYLOSGYAFATMGELIITLPLFSNRSVPLKTYIFAILLGA
LLAYLFFSDFSTIPNTSTTFDPAYWVIRIRITDFLSDLLIIAFAFPVIMVKIA
GSLYIVVYGSLFHSKIDPKAMYPPTGMFSVVCWFPLNTQNLIDFNRIKPIINVI
SLLLPLIYLIKSKALFGAKAH"
complement (5650.. .7065)
/codon_start=1
/transl_table=11
/product="yfla"
/protein_id="BAA22294.1"
/db_xref="GI:2443223"
/translation="MERLLVWIEHISDMLWGPPILIIILLTGLTYTILLKGFQFRYPL
YIFKQTISVGKKPGECTVTPLOALTSALSTIGAAINVGPAAIMFGPGGAVFWMW
LIALFAMAIFSESIVLAHVHREKNEQGVGGPMXYITKLRMKWLGVFFSVLIVEL
IPSNVQNSVSLSAETFSFNKIYAGIGIAFLIGLVVIGGVKRLGKYTEFVPLMAG
AYAGAGLILVLMNLSVPAFFSLVFSNAFTSSAVGGFAGAAETVRWGARGLYSN
EAGMGTAPIAHAAMTDHPVRQGPWSVIGVIDTLIICTTTAFIVLASGVMTGKNAS
DPAALTAAFOHYFGSGGGYFVSVSLVFVWSTIMVIFYGVKQAEFLFGRLAGHVIK
FYLAAILIGAAGGAKAIGVLDLALVIVPNVITALLLSRKVALYTEFTSQQY
LKDIRTKQKVPYPTKEAKNS"
complement (7063.. .7077)
complement (7149.. .7175)
complement (7190.. .7450)
/codon_start=1
/transl_table=11
/product="yflB"
/protein_id="BAA22295.1"
/db_xref="GI:2443224"
/translation="MGYRILKNAGMADEGALKKELMTIDHIEKYCEKEREQLIRK
KSEKQLLLDKLVKGMFSKPASAFYKKNVRYDLGRNRPSS"
complement (7432.. .7584)
/codon_start=1
/transl_table=11
/product="yflC"
/protein_id="BAA22296.1"
/db_xref="GI:2443225"
/translation="WMDFSHIVSEDKIKRAIKDGDFOQLPGMGKPLPKMTRHTRCRNC

```

```

AWAIVP"
complement (7448.. .7466)
complement (7582.. .7596)
complement (7600.. .7737)
/codon_start=1
/transl_table=11
/product="yflD"
/protein_id="BAA22297.1"
/db_xref="GI:2443226"
/translation="MREKAVNAVESTFRFDITKCKTRYLSRNKGIKWYIENCMIKYK
V"
complement (7735.. .7752)
7824.. .7840
7838.. .9787
/codon_start=1
/transl_table=11
/product="yflE"
/protein_id="BAA22298.1"
/db_xref="GI:2443227"
/translation="MKTIFRERGLAFFLIANVLMTKYVGVVLNENLGDNTIQKIL
LFVNPLSSSLFFLGLFKKLQQTALIIHFLMSFLLYANIYYRFRNDFITIPVI
MOAKTNGQLGDSAFSLMRPTDAFYFIDTIIILAIKVNKPAETSXKSFRIIPASS
ILVFLINLAVERSDRPELLTRSPDRNLYKLYLTNTYTIYDAVQNIKSNQRALADSS
DYTEVENYKANYDVNNVYFGKAEKNVYVLSLSQSPIDYKIDGKEVTPFLNKL
ADNETFYDFNFHQTCQKTSDAEFMENSILYPLAQSVFVNKAQNTLQSVFPAILKS
KNYTSATFHGTQTFWNRNEMYKAGIDKFFDSAYYDMNEENTKNYGMKDKPFKESM
PLLESIPQFPYTKFITSLSNHPFGMDGDDTDPAGDFGDSVDNYTQSAHYLDQSLIQ
FFNDLKDGILYDKSIIVMGDHYGI SENHNKAMAKVLGKDEITDYDNAQYLRVPLFIH
AAGVKGKHVHYAGDVAPTIHLGLVDTKDYLMGSDTILSKHEREVIFPRNGDFIS
PKYTKISGYDPTKTKELDESVDKSDSLVKKELEMSDKIINGDILLRYEPKGFKK
VNFSDYDTHKSDSDSETSKNEDK"
complement (9789.. .9820)
complement (9820.. .11178)
/codon_start=1
/transl_table=11
/product="yflF"
/protein_id="BAA22299.1"
/db_xref="GI:2443228"
/translation="MJSFLQKLGKSMPLPIAVLPVGLIILALGREDVFNIPFVYQAGT
AVFDPLIIFAIGIAIGISQKNGAAGLSAISYMLDAATKTIDKTNMNAVFGGIIA
GLIANGYVNFKOTKLEPGLFFSGRRLVPIITIIILAGIFGVVWPPIQSCINSF
GEMWLGIGGTGAGTIGLFLNRLIIPGLHVLNNIIFWFGYNGVTGDLARFPAKDP
AGWYTGFFPIMFGLPAACLAMVVTAKPSRKATAGMMIGFALTAFITGITETIEFA
FMFLSPLLYAVHAVLITGLSLFTVNWLGIRSGFSAGNIDYVLSYGAEKPLLLLVG
ICVAAVYFIVFYLIKALNKLTPOGDEDDDDVDLENTVDQVQNNENIMLKGIGKENLQ
TIDHCATRLRLTVKDTALVDEALKKAGAKGVKSGGSGVQVVIIGPNVFEAABELRAA
VK"
complement (11176.. .11195)
11334.. .11352
11350.. .12099
/codon_start=1
/transl_table=11
/product="yflG"
/protein_id="BAA22300.1"
/db_xref="GI:2443229"
/translation="MIVTNDQELGLKKIGRIVALAREMKRKAEPGGMSTKOLDLIGK
AVLDEHGAVSAPEKEYDPDPGTICISVNDVAHGIPSTSKILKAGDLVNIDISAEPPGF
YSDTGISFVLGEGBERHLKQCAENAFQKGLQQAQRQNRQIGRAVTHIARSQGT
VKTITGHGIGRISLHEAPNHIMNYDYPDNALFKNGTVIALEPFIETKAEITVEAGDG
WTFKTPDKSMVAQVEHTIVITKDEPIILTKL"
12166.. .12183
12181.. .12495
/codon_start=1

```

Alignment Scores:

Pred. No.:	2,09e-40	Length:	37900
Score:	679.50	Matches:	149
Percent Similarity:	62.78%	Conservatives:	50
Best Local Similarity:	47.00%	Mismatches:	103
Query Match:	40.21%	Indels:	15
DB:	1	Gaps:	4

US-10-724-972a-6352 (1-335) x D86417 (1-37900)

QY 8 LeuLysIleLeuSerValIleGlyLeuLeuPheValLeuIleAlaThrAlaLaCySGly 27
 Db 25123 ATTTGCGCATAGTGTGTTTAA---TTGCTCGCTGCTCATTTGATCCGCGTTCATCA 25179
 QY 28 AsnAsnSerSerSerAsnSerSerLysGluSerSerLysAspGlyValGluIleLysHis 47
 Db 25180 AGCCAGANTAAACCGAAGCGCAAGCGAGTCTAAGAGATTCCAGAGTGATCCATGAC 25239
 QY 48 GluGluGlyThrThrLysValProLysHisProLysArgValValValLysGluTyrSer 67
 Db 25240 GAAGAAGGAAACGACGACGAGTAAAGCGGCTGCTGCTGCTGCTGCTGCTGCTTCA 25299
 QY 68 PheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsnLys 87
 Db 25300 TTCTTGATGCGCTTCACAACTTCGCGATTACCGCGTGGCATCGCAGATGACACAAA 25359
 QY 88 LysAsnArgIleIleLysProLysArgAspLysIleGlyLysTyrThrSerValGlyThr 107
 Db 25360 AAAGATATGATTAAGAGCTTGCGCAGCTCATT---GATTACACATCTGTAGGCACA 25416
 QY 108 ArgLysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIleAlaAsp 127
 Db 25417 CGCAGCGAACCAATCTTCAGGTTCATCTCTTGAAGCTGATTAATCATCGCTGAC 25476
 QY 128 AsnAsnArgHisGlyGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeu 147
 Db 25477 GCTGAGCGCCATAAAACATTTATAACAGCTGAAAAAATCGCCCGACGATGAATTA 25536
 QY 148 LysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAla 167
 Db 25537 AAACCGGTGAAGCGACATATACGAAACGATCGACAGCTTTACGACCATCTGCTAAAGCA 25596
 QY 168 LeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGlu 187
 Db 25597 TTAATAAAGAGATGAAGGAAGAAAGAAAGCTTCCGAGCACAAGAAAGTCAATCAACG 25656
 QY 188 TyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLys 207
 Db 25657 CTAAAGCGCAACTTCGAAAGATGAAACCGCAACATCGTTCTCGCGTTCGAGAGCG 25716
 QY 208 SerGlyLeuAlaHisProSerAsnSerTyrValGlyGlnPheLysSerGlnLys 227
 Db 25717 GATTCCTTCAGCTTCATCATCATCTCTATGATGCGGAGAAATCTTTAAATGCTAGGC 25776
 QY 228 PheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLysGlyPro 247
 Db 25777 TTTACACAGCTGTGAGTCAAGTCAAGT-----AACGCC 25806
 QY 248 TyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMet 267
 Db 25807 TATCAAGAGGTCAGCTTGAGCAATTGAGCAAAATCGATCTGATATTTTCTTCTATCTCA 25866
 QY 268 ThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLysAspProValTyr 287
 Db 25867 GCCAACGAAGCAAAACC-----ATTGTAGATGATGTCGAAACGAAACCGCTCTGG 25917
 QY 288 LysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTyrAla 307
 Db 25918 AAAAATCTCAAGCGGTGAGAAATGGAAGATCTATGATCGGAGCGCTGACACTTGGACA 25977
 QY 308 ArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGlu 324
 Db 25978 AGATTACAGGATCAAGTCTAGTGAACAGCGCCAAAGATGTGCTTAA 26028
 RESULT 18
 BSUB0005/c 197409 bp DNA linear BCT 07-JUL-2003
 LOCUS Bacillus subtilis complete genome (section 5 of 21): from 813670 to 1011078.
 DEFINITION
 ACCESSION 299108 AL009126
 VERSION 299108.2 GI:32468715
 KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Bacillus subtilis subsp. subtilis str. 168
 Bacillus subtilis subsp. subtilis str. 168
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 1 (bases 1 to 197409)

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., Bertero, M.G., Bessieres, P., Bolotin, A., Borchert, S., Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S.C., Bron, S., Brouillet, S., Bruschi, C.V., Caldwell, B., Capuano, V., Carter, N.M., Choi, S.K., Codani, J.J., Conner, I.F., Cummings, N.J., Daniel, R.A., Denizot, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D., Emerson, P.T., Entian, K.D., Errington, J., Fabret, C., Ferrari, E., Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A., Galleron, N., Ghim, S.Y., Glaeser, P., Goffeau, A., Golligly, E.J., Grandi, G., Guisepi, G., Guy, B.J., Haga, K., Haiech, J., Harwood, C.R., Henaut, A., Hibert, H., Holbappel, S., Hosono, S., Hulio, M.F., Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y., Kladavich, B., Klen, C., Kobayashi, Y., Koetter, P., Konings, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A., Lardinois, S., Lauber, J., Lazarevic, V., Lee, S.M., Levine, A., Masuda, S., Mauch, C., Medigue, C., Medina, N., Mellado, R.P., Mizuno, M., Mosti, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M., Ogawa, K., Ogasawara, N., Oudega, B., Park, S.H., Parro, V., Pohl, T.M., Portetle, D., Rapoport, G., Rey, M., Reynolds, S., Rieger, M., Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadaie, Y., Sato, T., Scanlan, E., Schlaich, S., Schroeter, R., Scoffone, F., Sekiguchi, J., Sekowska, A., Seror, J.J., Serrero, P., Shin, B.S., Soldo, B., Sorokin, A., Tacconi, E., Takagi, T., Takahashi, H., Takemaru, K., Takeuchi, M., Tamakoshi, A., Tanaka, T., Terpstra, P., Tognoni, A., Tosato, V., Uchiyama, S., Vandenbol, M., Vannier, F., Vassarotti, A., Viari, A., Wambutt, R., Wedler, E., Wedler, H., Weizenegger, T., Winters, P., Wipat, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K., Yoshida, K., Yoshikawa, H.F., Zumbstein, E., Yoshikawa, H. and Danchin, A.

TITLE

Bacillus subtilis

NATURE 390 (6657), 249-256 (1997)

98044033

9384377

2 (bases 1 to 197409)

Kunst, F., Ogasawara, N., Yoshikawa, H. and Danchin, A.

Submitted (27-JUN-2003)

Genetique des Genomes Bacteriens, 28 rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr, adanchin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45 68 89 48

On Jul 7, 2003 this sequence version replaced gi:2633055.

This entry contains data from release R16.1 of the Subtilist database. Further data on gene annotation and detailed information about changes from previous releases can be found at

http://genolist.pasteur.fr/Subtilist/.

Location/Qualifiers

1..197409

/organism="Bacillus subtilis subsp. subtilis str. 168"

/mol_type="genomic DNA"

/strain="168"

/db_xref="taxon:224308"

complement(52..1608)

/genes="yfmM"

/locus_tag="BSU07420"

complement(52..1608)

/genes="yfmM"

/locus_tag="BSU07420"

/function="unknown"

/notes="similar to ABC transporter (ATP-binding protein)"

/codon_start=1

/transl_table=11

/protein_id="CAB12571.1"

/db_xref="GI:2633066"

/db_xref="GOA:O34512"

/db_xref="UniProt/TrEMBL:O34512"

/translation="MSILSVKLDLHGFGDRAIFNNVSRLLKGEHVLGANGEGKST"

FEATURES

source

gene

CDS

/function="unknown"		
/note="similar to macrolide-efflux transporter"		
/codon_start=1		
/transl_table=11		
/protein_id="CAB12575.1"		
/db_xref="GI:2633070"		
/db_xref="UniProt/TREMBL:O34440"		
/translation="MSKGYSLSSQTTHNGSLTMVVISFLFKMNSTTIASMTLVF LISIIPIRFGSAILPLIRIRUKLPITIIISQIQLMALLLCLFVALLRTYSNOTLILAV IISICISFNGWSPFLSKSAIGETIISPDRRKANGLLSTVDQTFQFVGWSLGGILIAF LQEGTIIITLFTLFTLSLLSLFCLPHGHANSVIREKNSPIVSGMWYLFQSKKLRL TIIINDLIESWAGMIWIGSVLSLAFVNEVLHKGESWMGFNGAYVLSGMIGGFIYKLS ERFQNKLNPMILIGAVSGSLTLVIGFTSNYSALAILVLPMGPAYILRLDLTQETLION ITEQTRINMSARSLSVQFIEMPSILAIGAISDFLGVRLVTVYSAGILLVLSAIIYGFS QLQFKKKKKHISF"		
gene	6397..6468	
/gene="yfmH"		
/locus_tag="BSU07470"		
CDS	6397..6468	
/gene="yfmH"		
/locus_tag="BSU07470"		
/function="unknown"		
/codon_start=1		
/transl_table=11		
/protein_id="CAB12576.1"		
/db_xref="GI:2633071"		
/db_xref="UniProt/TREMBL:O34890"		
/translation="MRKLRKNRKSIVSPVLITNEKW"		
gene	6535..7998	
/gene="yfmG"		
/locus_tag="BSU07480"		
CDS	6535..7998	
/gene="yfmG"		
/locus_tag="BSU07480"		
/function="unknown"		
/codon_start=1		
/transl_table=11		
/protein_id="CAB12577.1"		
/db_xref="GI:2633072"		
/db_xref="GOA:O34722"		
/db_xref="UniProt/TREMBL:O34722"		
/translation="MTLLMQDHKSINYKDIEILIEVEELITSVPENKKNDFKSLINDP IMTLNKHQHPVAENVFGWIACTPNTYIRYRLIDFVEKWPVYFSAVETIINLTQDP DLVSGRMDVCANHKTIEESVAVLSIIIDVRESSISYKPKVGLGAKVLSTLLDIFG EKHEELVNLKNIYDQNGILYLPNNDFPEKIPQSLIREFEKTEDGMLLIPGGPFIEGLA ENEPIDKPTKMDAVPRQKVLPPFFIDYFAFAYARKGKSLPTFTFQWEKAARGSKGNV NKQHRRTYWDRLYDHNHPVTGIDYFAFAYARKGKSLPTFTFQWEKAARGSKGNV WGDKFDPAKVQYAGSLYNEPITSLKSWRENLLKAHADKSLNHLTSDIPEQNGESPYG		
Alignment Scores:		
Pred. No.:	1.51e-39	Length: 197409
Score:	679.50	Matches: 149
Percent Similarity:	62.78%	Conservative: 50
Best Local Similarity:	47.00%	Mismatches: 103
Query Match:	40.21%	Indels: 15
DB:	1	Gaps: 4
US-10-724-972A-6352 (1-335) x BSUB0005 (1-197409)		
Qy	8	LeuYsilleLeuSerValIleGlyLeuLeuPheValLeuLeuIleAlaThrAlaIleCysGly 27
Dd	12378	ATTGCATCATGAGTGTTTAA--TTGTCGCCTGCCTCATTTGATCCGGCTGTTCA 1233
Qy	28	AnAanSerSerSerAanSerSerIysGluSerSerIysAspGlyValGluIleLysHis 47
Dd	12321	AGCCAGATAACACAGGAAGCGGCAAAAGCAGGCTTAAGGATTCGAGGTGATCCATGAC 1222
Qy	48	GluGluGlyThrThrLysValprolyHisProlySargValValValLeuGluIuTySer 67
Dd	12261	GAAGAAAGGAAAAACACACAGTAGCAGCGGCACACTAAGCGGGTGGTGTGCTTGAGCTTCA 1222
Qy	68	PheValAspAlaLeuValAlaLeuAspValIysProValGlyIleAlaAspAspAanLys 87

gene FNMIITGKLEPDEGKVENSKNVRGVGLDQHTVLEKGSIRDLVKDAFHFLFAMEEEMN
 EYINMGADPDELEKLLIEVGVIOQATNNNDVVDISKVEIARGLGLSDIGLERDV
 TDSGGQRTKVLVALLKLELLPGLDSEPTNYLDEQHI EWLRKYLQEVENAFILISHD
 IIFLNSVINLIYHVENQELTRYGDYHGFMEVIEYVKQQLAAAYKKQOQEVAEVKDFV
 ANKARVSRPNMAMSRQKIALKQMDI ELAAEKPKPEFHFHPARTSGKLIIFETKDLVIG
 VYDPSLRPLNRMERGOKIALYKAKTITKLLSGLEIQPLSGSVRGEHIYTGVF
 EQEVETNNNTCIEEWSFSPSYTOVEIRAAAPAKCGLTTHIEISRVSVLSGGEKAKVR
 LCKLINSETNLLVLEPTNHLDDADAKEELKALKEYSIKVSGSILLISHEPDPFYMDIATETW
 NCESMTTKVL"
 1781..2911
 /gene="yfmL"
 /locus_tag="BSU07430"
CDS 1781..2911
 /gene="yfmL"
 /locus_tag="BSU07430"
 /function="unknown"
 /note="similar to RNA helicase"
 /codon_start=1
 /transl_table=11
 /protein_id="CAB12572.1"
 /db_xref="GI:2633067"
 /db_xref="GOA:O34750"
 /db_xref="UniProt/TREMBL:O34750"
 /translations="MTQTPWFLHNAQSFIOENNNASGFOKPTVQEQAAQLIMDGKDV
 IASPTGTGLTAYALPVERIEIKPEQKQPAVILASRELVMQIFQVTDWKAGASLR
 AASLIGANVKKQVEXLKKHPHITVGTGPRVFELIKAKLKMHEVKTIVLDETDLVL
 PEHRTMKLIKTILDRQQLCFSAI LKKTEDVLRELAQEPEVLKQVRSKAEAGVKV
 HQYLICDRQKLLQKSLGKQALVFRDIGNLSVYAEKLAHVHVELGVLHSEAK
 KMERAKIATFDEGFPELLATDIAARGLDIENLPYIHADIPDEDGYVHRSGRGTGRA
 KGEKNVLSVLTKLEESKLKKMAKKLVGELSEAVYAGGKLTKT"
 2979..3425
 /gene="yfmK"
 /locus_tag="BSU07440"
 2979..3425
 /gene="yfmK"
 /locus_tag="BSU07440"
 /function="unknown"
 /codon_start=1
 /transl_table=11
 /protein_id="CAB12573.1"
 /db_xref="GI:2633068"
 /db_xref="GOA:O34536"
 /db_xref="UniProt/TREMBL:O34536"
 /translations="MASIDRFQMQPEPKILSHWFENEDTRRRMDGMLLLDEWYARV
 NKDHDVIMAYDQLPAGMVVIFSGERTYIGLIVNPLVRLKGYGKQILQLKLTWTEPD
 FTSVREWACILEENRSLACFQAAGFTLEDTEPDEDFLTILRN"
 complement(3478..4497)
 /gene="yfmJ"
 /locus_tag="BSU07450"
 complement(3478..4497)
 /gene="yfmJ"
 /locus_tag="BSU07450"
 /function="unknown"
 /note="similar to quinone oxidoreductase"
 /codon_start=1
 /transl_table=11
 /protein_id="CAB12574.1"
 /db_xref="GI:2633069"
 /db_xref="GOA:O34812"
 /db_xref="UniProt/TREMBL:O34812"
 /translations="MTASQQQTQLARRPQGI PVHEDPRFETIPVPEKQGEVLVKTLYL
 VSDVPMGRMDQTKSVYEPFALDKALSGSVIAEVVSDGNHLKKGDIVIGNLSWQEFES
 AVSESALRKIDTSLPASAYALIGLQMTGLTAYFGLLDIGRPKEGTIVVSGAAGVGS
 VYSDALRKIDTSLPASAYALIGLQMTGLTAYFGLLDIGRPKEGTIVVSGAAGVGS
 TVGQIAKIGARVGIAGSDEKIDYKQELQFDEAINYKTADDIQKALQNAQCPGVVDV
 YFDVYGGPISDAVMNLLNEFARI PVCGAISYNASEADDMPRVQSKLIKTKSLMQG
 FTVSDYSIDRFSEGAQLAEWLKAGKLHYEETITEGFENIPDAFLGLFKGENKQGLINL
 VSDPS"
 complement(4979..6199)
gene /gene="yfmI"
 /locus_tag="BSU07460"
CDS complement(4979..6199)
 /gene="yfmI"
 /locus_tag="BSU07460"
 complement(4979..6199)

Db	12201	TTCTTGGATCGGTTCAATCTCGGCATTACGCGGTGGGCATCGCAGATGACAACAAA	12142
Qy	88	LysAsnArgIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThr	107
Db	12141	AAAGATATGATTAATAAGAGCTGTGCGGACGCTCAATT--GATTACATCTGTAGGCACA	12085
Qy	108	ArgLysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIleAlaAsp	127
Db	12084	CGCAGCGAACCACCACTCTTGAGGTTCATCAGTTCCTTGAAGCTGATTAAATCATCGCTGAC	12025
Qy	128	AsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeu	147
Db	12024	GCTGAGCGCCATATAAACAATTTATTAACAGCTGAAAAAATTCGCCGCGAGTGAATTA	11965
Qy	148	LysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAla	167
Db	11964	AAAAAGCGTCAGGACGATATATGACGAACCATGACGACGCTTACGACCATGCTGAAGCA	11905
Qy	168	LeuGlyLysGluGluGluGlyLysLysArgLysGluGluHisAspLysIleGluGlu	187
Db	11904	TTAAATAAAGAAAGATGAAGAAAGAAAGCTTCCCGGACACAAAAAGTCATCAACGAT	11845
Qy	188	TyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLys	207
Db	11844	CTAAAGCCGNACTTCGAAAGATGAACCCGCAACATCGTCTCGCGGTGCAAGCGC	11785
Qy	208	SerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGly	227
Db	11784	GATTCCTTCAGCTTCATACATCATCATCTATGACGAGAAATCTTTAAATGCTAGGC	11725
Qy	228	PheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGlyPro	247
Db	11724	TTTACACACGCTGTGAAGTCAGAT-----	11695
Qy	248	TyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheLeuMet	267
Db	11694	TATCAAGAGTCAAGCTTCAGCAATTCGACAAATCGATCTGATATTTTGTTCATCTCA	11635
Qy	268	ThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLysAspProValTyr	287
Db	11634	GCCACGAGGCAAAACC-----ATTGTAGATGATGTGAAACGAAACCCGCTCTGG	11584
Qy	288	LysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTyrAla	307
Db	11583	AAAAATCTCAAGCGGTGAAAAATGGAACAGTCTATGATGCGGACCGTGCACATTTGGACA	11524
Qy	308	ArgSerArgGlyLeuIleSerSerGluMetAlaLysGluLeuValGlu	324
Db	11523	AGATTCAGAGGCATCAAGTCTAGTGAACAGCGCCAAAGATGTGCTTAAA	11473
RESULT 19			
BSY14084			
LOCUS			
DEFINITION	BSY14084 Bacillus subtilis chromosomal DNA, region 78-80 degrees: aprE to comK.		
ACCESSION	Y14084		
VERSION	Y14084.1 GI:2226250		
KEYWORDS	comK gene; ynfO gene; ynfP gene; ynfQ gene; ynfR gene; ynfS gene; ynfT gene; ynfU gene; ynfV gene; ynfW gene; ynfX gene; ynfY gene.		
SOURCE	Bacillus subtilis		
ORGANISM	Bacillus subtilis		
REFERENCE	1 (bases 1 to 104)		
AUTHORS	Stahl, M.L. and Ferrari, E.		
TITLE	Replacement of the Bacillus subtilis subtilisin structural gene with an in vitro-derived deletion mutation		
JOURNAL	J. Bacteriol. 158 (2), 411-418 (1984)		
MEDLINE	84212198		
PUBMED	6427178		
REFERENCE	2 (bases 1 to 441)		
AUTHORS	Park, S.S., Wong, S.L., Wang, L.F. and Doi, R.H.		
TITLE	Bacillus subtilis subtilisin gene (aprE) is expressed from a sigma A (sigma 43) promoter in vitro and in vivo		
JOURNAL	J. Bacteriol. 171 (5), 2657-2665 (1989)		
MEDLINE	89213955		
PUBMED	2496113		
REFERENCE	3 (bases 10610 to 12556)		
AUTHORS	van Sinderen, D., ten Berge, A., Hayema, B.J., Hamoen, L. and Venema, G.		
TITLE	Molecular cloning and sequence of comK, a gene required for genetic competence in Bacillus subtilis		
JOURNAL	Mol. Microbiol. 11 (4), 695-703 (1994)		
MEDLINE	94254725		
PUBMED	8196543		
REFERENCE	4 (bases 1 to 12556)		
AUTHORS	Noback, M.A., Terpstra, P., Holsappel, S., Venema, G. and Bron, S.		
TITLE	Direct Submission		
JOURNAL	Submitted (26-JUN-1997) Noback M. A., University of Groningen, Dept. of Genetics, Kerklaan 30, Haren, 9751 NN The Netherlands		
FEATURES	Location/Qualifiers		
source	1..12556		
	/organism="Bacillus subtilis"		
	/mol_type="genomic DNA"		
	/strain="168"		
	/db_xref="taxon:1423"		
	/map="78 to 80 degrees, aprE to comK"		
gene	383..852		
	/gene="yhfO"		
RBS	383..396		
	/gene="yhfO"		
CDS	403..852		
	/gene="yhfO"		
	/note="Similarity to several hypothetical proteins, from Haemophilus influenzae (Swiss Prot P4036), Saccharomyces cerevisiae (Swiss Prot P39979), and others."		
	/citation=[2]		
	/codon_start=1		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="CAA74538.1"		
	/db_xref="GI:2226251"		
	/db_xref="GOA:O07614"		
	/db_xref="UniProt/TrEMBL:O07614"		
	/translation="MNIRQAQTSDAIAIPLFNQVREYFVROASDIQGAFAFLKARLEN HESVILIAEENGFEFTGLYPTSSVSMKRIYILNDLVVPHARTKAGGRLLSAAK DYAGONGAKCLTLQTEHNRKARSLEYQNGYEEDTGFVHYCLINLPAK"		
	907..1916		
gene	/gene="yhfP"		
RBS	907..919		
	/gene="yhfP"		
CDS	924..1916		
	/gene="yhfP"		
	/note="high similarity to yhdH from E.coli (Swiss Prot P26846), and lower to many alcohol dehydrogenases."		
	/codon_start=1		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="CAA74539.1"		
	/db_xref="GI:2226252"		
	/db_xref="GOA:O07615"		
	/db_xref="UniProt/TrEMBL:O07615"		
	/translation="MSTLFOALQAEKNADDVSVHVKTITSTEDLPKGVLIKVAYSGIN YKDLGKAGKGNIVREYPLILIGIDAQCTVSSNDPRAEGDEVATSYELGVSDDGGL SEVASVPQDWLPLPQNLSEKAMVYGTAGTAAALSVRLEQNGLSPEKGSVLVTGAT CGVGGIASMLNKRGYDVASTGNREAADYIKQLGASEVTSREDYVTLRGLKALSKQOW GAVDPVVGKQLASLSKIKYGGSAVSGVLTGGEVPATYVFLRGLVSLIGDSVYC PMDVRAAWRMSSDLKPDQLLTIVDREVSLEETPGALKDILQNRIGRVIVKL"		
	2043..3104		
gene	/gene="yhfQ"		
RBS	2043..2056		
	/gene="yhfQ"		
CDS	2058..3104		
	/gene="yhfQ"		
	/note="Similarity to citrate-dependent iron transport protein fecB from E.coli (PIR S56515) and Synecchocystis sp. (EMBL: D90899)."		
	/codon_start=1		

```

/transl_table=11
/product="hypothetical protein"
/protein_id="CAA74540.1"
/db_xref="GI:2226253"
/db_xref="GOA:O07616"
/db_xref="UniProt/TREMBL:O07616"
/translation="MILIPCVTTGLNHRNYQOERLEHMKTKLILTVLLLSVLTAACSSSGNQNSKEHKVAVTHDLGKTNVPEPKRVVLELGFIDTLDLGLTTPGVADDAKAKILINKLKKIKDGTSGVTSQFSQSMEKIASLKDPLDIADTTSHKKYVDOLKKIAPTIALNNADYQDTIDASLTAKAVGKEKEMKKLTAHEEKLSQTKKLSANSQSVLLIGTNDTIRMARDENFTSRLLTQVGYAIYSTGSDSSNGDGSVNMKMTLQLLKTTDVLILMTGKTDDLADGKRPKEKNVLKKLKVAKNGHVYHVDRAVMSLRSDVGANAIDLQEKEMPAKK"
3099..3141
terminator
gene
complement(3136..3731)
/gene="yhfr"
CDS
complement(3136..3717)
/gene="yhfr"
/notes="similarity to a hypothetical protein o215b from E.coli (Swiss Prot P36942), and to several phosphoglycerate mutase (pgm) proteins, from Treponema pallidum (EMBL: U55214) and Schizosaccharomyces pombe (Swiss Prot P36623)."
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAA74541.1"
/db_xref="GI:2226254"
/db_xref="GOA:O07617"
/db_xref="UniProt/TREMBL:O07617"
/translation="MTAVCLVHGSETDNLQCKQKGTIDPLNATGEROARETGEYVKDFSWDIIVTSPUKRAKRTPIINLYHLPIVEMDDPKERDYGDAEGMPLERTKRPYDNTFMYNETLEELTDLRMGLAKVQNPVKKVLIVAHGAAIHALLTEISGGDFELQSTRVUNACLGNIEFAEEKRIKOYNTNSHLSGFK"
complement(3722..3731)
RBS
/gene="yhfr"
complement(3788..4882)
/gene="yhfr"
complement(3788..4882)
/gene="yhfr"
/notes="similarity to acetyl-CoA acetyltransferases (ThiL, encoding the enzyme for the first step in poly-beta-hydroxybutyrate synthesis) from Chromatium vinosum (Swiss Prot P45369), Alcaligenes eutrophus (Swiss Prot P14611), Thio cystis violacea (Swiss Prot P45363), and B. subtilis (Swiss Prot P45855)."
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAA74542.1"
/db_xref="GI:2226255"
/db_xref="UniProt/TREMBL:O07618"
/translation="MNAVIVDAKRTIFGNQGLLKKPFLDPEDLAAPITRCISRKLEQVQDEVILNATCGRGNLRSALQAGLPLSVPGMTIDRCQSGLEAVRACSLIQAGAGTMYTAGSSSSQPSFERARFPDAIGDPMGIAEYTAARYSISRSMDQYALLSHQSRNADHGFGEVEVVALGELETDRFAFLKTRIEPILIPRAKVPDTSSTGTVTAANSQIADGAAALVMESEKALGLKPFVIFGSAVSGHPNPAPVPAIRQLLHTHDVTDDIDLFPEINAEPAVKIVCSQELGIPGKSNVRGGALALGHPYGASGAALVTRLFVEAKRRPDCOVAAVAGSGGIGLALLFEVLA"
complement(4879..6318)
gene
/gene="yhfr"
complement(4879..6318)
CDS
/gene="yhfr"
/notes="similarity to long-chain-acyl-CoA synthetase from B. subtilis (EMBL: 275208), bile acid-CoA ligase from Eubacterium sp. (Swiss Prot P19409), long-chain-fatty acid:CoA ligase from E.coli (Lcfa; Swiss Prot P29212)."
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAA74543.1"
/db_xref="GI:2226256"
/db_xref="GOA:O07618"

```

Db	2550	AATAATTTAAATGCTGATTATCAGGATACATTCAGCGCTTCGCTTACGATTGCAAAAGCA	2609
Qy	168	LeuGlyLysGluGluGluGlyLysLysArgLeuGluGluHisAspLysLysLeuGluGlu	187
Db	2610	GTCGCGCAAGGAGAAGAAATCGAGAAAAGCTGACGCGCATGAAGAAAAGCTTAGCGAG	2669
Qy	188	TyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLys	207
Db	2670	ACAAAGCAGAAAATCAGCGCGAAC--AGCCAGTCCGTGCTTTTGATCGGAAATACAAAT	2726
Qy	208	SerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGly	227
Db	2727	GATACCATTTATGGCCAGGAGTGAATACTTCTTTACATCGAGACTTTTAAACACAGGTGGC	2786
Qy	228	PheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGly---	246
Db	2787	TACCGATATGCAATTCAGT-----ACGTGAGGCAATAGCGATTCCAACCAATGGCGGC	2837
Qy	247	-----ProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluAArgMet	264
Db	2838	GACTCTGTGAATATGAAAATGACACTGGAGCAGCTGTGTGAATAACAGATCCGGATGTGATC	2897
Qy	265	PheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGlu---LeuGluLys	283
Db	2898	ATCCTGTATGACAGGAAAAAACAGATGACCTCGAGCGCCAGCGTAAACCCCGATCGAAAAG	2957
Qy	284	AspProValTyrLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg	303
Db	2958	AATGTCCTTTTGGAAAAAATGAAAGCGAGTGAATAAACCGGCATGTATACACAGGTGGATCGT	3017
Qy	304	AspLeuTrpAlaArgSerArgGlyLeuIleSerSer-----GluGluMetAla	319
Db	3018	CGCGTGTGTCTCTCGCCCGCAGTGATAGCGGGCGGAATGCCATTTTGGACGAGCTTCAA	3077
Qy	320	LysGluLeuValGluLeuSerLysLysAspSerLysLysAspAsn	334
Db	3078	AAGAGATCGCGCTGCTAGAAATATAAGAAAAGACGAGCAAC	3122

RESULT 20

BSUB0006	198743 bp	DNA	linear	BCT 07-JUL-2003
LOCUS				
DEFINITION	Bacillus subtilis complete genome (section 6 of 21): from 1011039 to 1209781.			
ACCESSION	Z99109 AL009126			
VERSION	Z99109.2			
KEYWORDS	GI:32468723			
SOURCE	Bacillus subtilis subsp. subtilis str. 168			
ORGANISM	Bacillus subtilis subsp. subtilis str. 168			
REFERENCE	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
1 (bases 1 to 198743)				
AUTHORS	Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G.			

REFERENCE
1 (bases 1 to 198743)

AUTHORS

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., Azevedo, V., Bertero, M. G., Bessieres, P., Bolotin, A., Borchert, S., Borriass, R., Boursier, L., Brans, A., Braun, M., Brignell, S. C., Bron, S., Brouillet, S., Bruschi, C. V., Caldwell, B., Capuano, V., Carter, N. M., Choi, S. K., Codani, J. J., Connerth, I. F., Cummings, N. J., Daniel, R. A., Denizot, F., Devine, K. M., Dusterhof, A., Ehrlich, S. D., Emerson, P. T., Enriani, K. D., Errington, J., Fabret, C., Ferrari, E., Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Gallazzi, A., Galleron, N., Ghim, S. Y., Glaeser, P., Goffeau, A., Golithly, E. J., Grandi, G., Guseppi, G., Guy, B. J., Haga, K., Haiech, J., Harwood, C. R., Henaut, A., Hibert, H., Holsappel, S., Hosono, S., Hullo, M. F., Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y., Klaer-Blanchard, M., Klein, C., Kobayashi, Y., Koetter, P., Konigstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A., Lardinois, S., Lauber, J., Lazarevic, V., Lee, S. M., Levine, A., Liu, H., Masuda, S., Mauel, C., Medigue, C., Medina, N., Mellado, R. P., Mizuno, M., Mostl, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M., Ogawa, K., Ogiwara, A., Oudega, B., Park, S. H., Parro, V., Pohl, T. M., Portetelle, D., Porwollik, S., Prescott, A. M., Presecan, E., Pujic, P., Purnelle, B., Rapoport, G., Ray, M., Reynolds, S., Rieger, M., Rivalta, C., Rocha, E. S., Roche-B, Rose, M., Sadaie, Y., Sato, T., Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J.,

Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J.,

Sekowska, A., Seror, S.J., Seror, P., Shin, B.S., Soldo, B.,
 Sorokin, A., Tacconi, E., Takagi, T., Takahashi, H., Takemaru, K.,
 Takeuchi, M., Tamakoshi, A., Tanaka, T., Terpstra, P., Tognoni, A.,
 Tosato, V., Uchiyama, S., Vandenbol, M., Vannier, F., Vassarotti, A.,
 Viari, A., Wamboldt, R., Wedler, E., Wedler, H., Weitzengger, T.,
 Winters, P., Wipat, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K.,
 Yoshida, K., Yoshikawa, H.F., Zumstein, E., Yoshikawa, H. and
 Danchin, A.
 The complete genome sequence of the gram-positive bacterium
Bacillus subtilis
 Nature 390 (6657), 249-256 (1997)
 98044033
 9384377
 2 (bases 1 to 198743)
 Kunst, F., Ogaawara, N., Yoshikawa, H. and Danchin, A.
 Direct Submission
 Submitted (27-JUN-2003) I. Moszer, A. Danchin, Institut Pasteur,
 Genetique des Genomes Bacteriens, 28 rue du Docteur Roux, 75724
 Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr,
 adanchin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45
 68 89 48
 On Jul 7, 2003 this sequence version replaced gi:2633260.
 This entry contains data from release R16.1 of the Subtilist
 database. Further data on gene annotation and detailed information
 about changes from previous releases can be found at
<http://genolist.pasteur.fr/Subtilist/>.

FEATURES
source

```

/organism="Bacillus subtilis subsp. subtilis str. 168"
/mol_type="genomic DNA"
/strain="168"
/db_xref="taxon:224308"
/complement(38..1504)
/gene="lytF"
/locus_tag="BSU09370"
/complement(38..1504)
/gene="lytF"
/locus_tag="BSU09370"
/function="cell wall lytic activity"
/note="alternate gene name: cwle, yhdD"
/codon_start=1
/transl_table=11
/product="gamma-D-glutamate-meso-diaminopimelate
muropetidase (major autolysin) (CWBP49)"
/protein_id="CAB12776.1"
/db_xref="GI:2633272"
/db_xref="GOA:O07532"
/db_xref="UniProt/TrEMBL:O07532"
/translations="MKKKAAGLTASAIWGLTVTPAEATIKVKGDSLMLQAQTY
NTSVAALTSANHLSTTVLSIGQTLTIPGSKSTSSSTSTTKGSSVTVTKSGDSL
WLANEFKMTQBELKGLGDSLLIRAGQKLVGVSTVSSSSSSSKSNKSSSSSSSK
SSNKNKSSSSSSSTGTYYVGLGDSLWIAKNKVNMSIAELKVLNNLKSDTITVNNQVLTK
SSGSDTSKDNKSSKNQTSATTKYTVKSGDSLWKIANNYNLTQQIIRNNKSLDVL
YGVQVLKTKKSSSSSSSTGTVTKSGDSLWVI AQKNVTAQRIREK
NLIKTDVLOGQKLVISGKASSSSSSGNSTTSSTSAKINTWISAQAQLGVPYRWGG
TTPSGPDCSGFIYYVLNKVTSYSLTAAGYWNMTKSVQPAVGDVFFFTSTYKAGFSHV
GIYLNGEFINANDSGVWISNNMNSTWKORYLQAKRYF"
/complement(1657..2097)
/gene="yhdE"
/locus_tag="BSU09380"
/complement(1657..2097)

```

gene

```

/locus_tag="BSU09380"
complement(1657..2097)
/gene="yhxE"
/locus_tag="BSU09380"
/function="unknown"
/notes="similar to unknown proteins"
/codon_start=1
/transl_table=11
/protein_id="CAB12777.1"
/db_xref="GI:2633273"
/db_xref="GGA:007573"
/db_xref="UniProt/Swiss-Prot:O07573"
/translation="MKLTNYDYSLRVLFLAAERPGELSNIKQIAETYSISKHLMK
VIVRLGOLGVETIRGGGCIIRLGMDDPPDINTEVVRKTEDPDNTEVFEDAKMKI

```


Qy	208	SerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGly	227
Db	96593	GATACCATATTGGCCAGGGATGAATACTTTTACATCGAGACTTTTAACACAGGTGGC	96652
Qy	228	PheLysGluLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGly---	246
Db	96653	TACCGATATGCAATCAGT-----ACGTCAGGCAATAGCGATTCAAGCAATGGCGC	96703
Qy	247	-----ProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMet	264
Db	96704	GACTCTGTGAATAGAAATTCACACTGGAGCAGCTCTGAAAACAGATCCGGATGTGATC	96763
Qy	265	PheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGlu---LeuGluLys	283
Db	96764	ATCCTGATGCACGAAAAACAGATGACCTTCAGCGCGCGTAAACGCCCATCGAAAAG	96823
Qy	284	AspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg	303
Db	96824	AATGTCCTTTGGAAAAAACACTGAAGGCAGTGAAAAACGGGCATGTATACCACGGATCGT	96883
Qy	304	AspLeuTrpAlaArgSerArgGlyLeuIleSerSer-----GluGluMetAla	319
Db	96884	GCGGTGTGGTCTCTGCGCGCAGTGTAGACGGGCGCAATGCCATTTTGGACAGCTTCAA	96943
Qy	320	LysGluLeuValGluLeuSerLysLysAspSerLysLysAspAsn	334
Db	96944	AAAGAGATCCGCGTCTCTAAGAAATAAAGAAAAAGACAGGCAAAAC	96988

RESULT 21
AE017133 11

RESULT 21
AE017313 11

WPCOMMENT_14

Sequence split into 43 fragments LOCUS AE017722 1000000 bp AE017722

Fragment Name	Begin	End
Fragment 1	0000	0000
Fragment 2	0001	0001
Fragment 3	0002	0002
Fragment 4	0003	0003
Fragment 5	0004	0004
Fragment 6	0005	0005
Fragment 7	0006	0006
Fragment 8	0007	0007
Fragment 9	0008	0008
Fragment 10	0009	0009
Fragment 11	0010	0010
Fragment 12	0011	0011
Fragment 13	0012	0012
Fragment 14	0013	0013
Fragment 15	0014	0014
Fragment 16	0015	0015
Fragment 17	0016	0016
Fragment 18	0017	0017
Fragment 19	0018	0018
Fragment 20	0019	0019
Fragment 21	0020	0020
Fragment 22	0021	0021
Fragment 23	0022	0022
Fragment 24	0023	0023
Fragment 25	0024	0024
Fragment 26	0025	0025
Fragment 27	0026	0026
Fragment 28	0027	0027
Fragment 29	0028	0028
Fragment 30	0029	0029
Fragment 31	0030	0030
Fragment 32	0031	0031
Fragment 33	0032	0032
Fragment 34	0033	0033
Fragment 35	0034	0034
Fragment 36	0035	0035
Fragment 37	0036	0036
Fragment 38	0037	0037
Fragment 39	0038	0038
Fragment 40	0039	0039
Fragment 41	0040	0040
Fragment 42	0041	0041
Fragment 43	0042	0042
Fragment 44	0043	0043
Fragment 45	0044	0044
Fragment 46	0045	0045
Fragment 47	0046	0046
Fragment 48	0047	0047
Fragment 49	0048	0048
Fragment 50	0049	0049
Fragment 51	0050	0050
Fragment 52	0051	0051
Fragment 53	0052	0052
Fragment 54	0053	0053
Fragment 55	0054	0054
Fragment 56	0055	0055
Fragment 57	0056	0056
Fragment 58	0057	0057
Fragment 59	0058	0058
Fragment 60	0059	0059
Fragment 61	0060	0060
Fragment 62	0061	0061
Fragment 63	0062	0062
Fragment 64	0063	0063
Fragment 65	0064	0064
Fragment 66	0065	0065
Fragment 67	0066	0066
Fragment 68	0067	0067
Fragment 69	0068	0068
Fragment 70	0069	0069
Fragment 71	0070	0070
Fragment 72	0071	0071
Fragment 73	0072	0072
Fragment 74	0073	0073
Fragment 75	0074	0074
Fragment 76	0075	0075
Fragment 77	0076	0076
Fragment 78	0077	0077
Fragment 79	0078	0078
Fragment 80	0079	0079
Fragment 81	0080	0080
Fragment 82	0081	0081
Fragment 83	0082	0082
Fragment 84	0083	0083
Fragment 85	0084	0084
Fragment 86	0085	0085
Fragment 87	0086	0086
Fragment 88	0087	0087
Fragment 89	0088	0088
Fragment 90	0089	0089
Fragment 91	0090	0090
Fragment 92	0091	0091
Fragment 93	0092	0092
Fragment 94	0093	0093
Fragment 95	0094	0094
Fragment 96	0095	0095
Fragment 97	0096	0096
Fragment 98	0097	0097
Fragment 99	0098	0098
Fragment 100	0099	0099

AE017333_00	1	110000
-------------	---	--------

AE017333_01	100001	210000
-------------	--------	--------

AE017333_02	200001	310000
-------------	--------	--------

AE017333_03	300001	410000
AE017333_04	400002	500000

AE017333_04 400001 510000
AE017333_05 E00001 C10000

AE017333-06
600001
710000
610000

AE017333-07
700001
810000

AE017333_08
800001
910000

AE017333_09	900001	1010000
-------------	--------	---------

AE017333_10	1000001	1110000
-------------	---------	---------

AE017333_11	1100001	1210000
AE017333_12	1200002	1210001
AE017333_13	1200003	1210002
AE017333_14	1200004	1210003
AE017333_15	1200005	1210004
AE017333_16	1200006	1210005
AE017333_17	1200007	1210006
AE017333_18	1200008	1210007
AE017333_19	1200009	1210008
AE017333_20	1200010	1210009
AE017333_21	1200011	1210010
AE017333_22	1200012	1210011
AE017333_23	1200013	1210012
AE017333_24	1200014	1210013
AE017333_25	1200015	1210014
AE017333_26	1200016	1210015
AE017333_27	1200017	1210016
AE017333_28	1200018	1210017
AE017333_29	1200019	1210018
AE017333_30	1200020	1210019
AE017333_31	1200021	1210020
AE017333_32	1200022	1210021
AE017333_33	1200023	1210022
AE017333_34	1200024	1210023
AE017333_35	1200025	1210024
AE017333_36	1200026	1210025
AE017333_37	1200027	1210026
AE017333_38	1200028	1210027
AE017333_39	1200029	1210028
AE017333_40	1200030	1210029
AE017333_41	1200031	1210030
AE017333_42	1200032	1210031
AE017333_43	1200033	1210032
AE017333_44	1200034	1210033
AE017333_45	1200035	1210034
AE017333_46	1200036	1210035
AE017333_47	1200037	1210036
AE017333_48	1200038	1210037
AE017333_49	1200039	1210038
AE017333_50	1200040	1210039
AE017333_51	1200041	1210040
AE017333_52	1200042	1210041
AE017333_53	1200043	1210042
AE017333_54	1200044	1210043
AE017333_55	1200045	1210044
AE017333_56	1200046	1210045
AE017333_57	1200047	1210046
AE017333_58	1200048	1210047
AE017333_59	1200049	1210048
AE017333_60	1200050	1210049
AE017333_61	1200051	1210050
AE017333_62	1200052	1210051
AE017333_63	1200053	1210052
AE017333_64	1200054	1210053
AE017333_65	1200055	1210054
AE017333_66	1200056	1210055
AE017333_67	1200057	1210056
AE017333_68	1200058	1210057
AE017333_69	1200059	1210058
AE017333_70	1200060	1210059
AE017333_71	1200061	1210060
AE017333_72	1200062	1210061
AE017333_73	1200063	1210062
AE017333_74	1200064	1210063
AE017333_75	1200065	1210064
AE017333_76	1200066	1210065
AE017333_77	1200067	1210066
AE017333_78	1200068	1210067
AE017333_79	1200069	1210068
AE017333_80	1200070	1210069
AE017333_81	1200071	1210070
AE017333_82	1200072	1210071
AE017333_83	1200073	1210072
AE017333_84	1200074	1210073
AE017333_85	1200075	1210074
AE017333_86	1200076	1210075
AE017333_87	1200077	1210076
AE017333_88	1200078	1210077
AE017333_89	1200079	1210078
AE017333_90	1200080	1210079
AE017333_91	1200081	1210080
AE017333_92	1200082	1210081
AE017333_93	1200083	1210082
AE017333_		

AE017333_12	1200001	1310000
AE017333_13	1200002	1410000

AE017333-13	1300001	1410000
AE017333-14	1400001	1510000

AE017333 15
1500001 1610000
1500001 1500000

AE017333_16
1600001
1710000

AE017333_17	1700001	1810000
-------------	---------	---------

AE017333_18
1800001 · 1910000

AE017333_19	1900001	2010000
AE017333_20	2000001	2010000
AE017333_21	2000002	2010000
AE017333_22	2000003	2010000
AE017333_23	2000004	2010000
AE017333_24	2000005	2010000
AE017333_25	2000006	2010000
AE017333_26	2000007	2010000
AE017333_27	2000008	2010000
AE017333_28	2000009	2010000
AE017333_29	2000010	2010000
AE017333_30	2000011	2010000
AE017333_31	2000012	2010000
AE017333_32	2000013	2010000
AE017333_33	2000014	2010000
AE017333_34	2000015	2010000
AE017333_35	2000016	2010000
AE017333_36	2000017	2010000
AE017333_37	2000018	2010000
AE017333_38	2000019	2010000
AE017333_39	2000020	2010000
AE017333_40	2000021	2010000
AE017333_41	2000022	2010000
AE017333_42	2000023	2010000
AE017333_43	2000024	2010000
AE017333_44	2000025	2010000
AE017333_45	2000026	2010000
AE017333_46	2000027	2010000
AE017333_47	2000028	2010000
AE017333_48	2000029	2010000
AE017333_49	2000030	2010000
AE017333_50	2000031	2010000
AE017333_51	2000032	2010000
AE017333_52	2000033	2010000
AE017333_53	2000034	2010000
AE017333_54	2000035	2010000
AE017333_55	2000036	2010000
AE017333_56	2000037	2010000
AE017333_57	2000038	2010000
AE017333_58	2000039	2010000
AE017333_59	2000040	2010000
AE017333_60	2000041	2010000
AE017333_61	2000042	2010000
AE017333_62	2000043	2010000
AE017333_63	2000044	2010000
AE017333_64	2000045	2010000
AE017333_65	2000046	2010000
AE017333_66	2000047	2010000
AE017333_67	2000048	2010000
AE017333_68	2000049	2010000
AE017333_69	2000050	2010000
AE017333_70	2000051	2010000
AE017333_71	2000052	2010000
AE017333_72	2000053	2010000
AE017333_73	2000054	2010000
AE017333_74	2000055	2010000
AE017333_75	2000056	2010000
AE017333_76	2000057	2010000
AE017333_77	2000058	2010000
AE017333_78	2000059	2010000
AE017333_79	2000060	2010000
AE017333_80	2000061	2010000
AE017333_81	2000062	2010000
AE017333_82	2000063	2010000
AE017333_83	2000064	2010000
AE017333_84	2000065	2010000
AE017333_85	2000066	2010000
AE017333_86	2000067	2010000
AE017333_87	2000068	2010000
AE017333_88	2000069	2010000
AE017333_89	2000070	2010000
AE017333_90	2000071	2010000
AE017333_91	2000072	2010000
AE017333_92	2000073	2010000
AE017333_93	2000074	2010000
AE017333_94	2000075	2010000
AE017333_95	2000076	2010000
AE017333_96	2000077	2010000
AE017333_97	2000078	2010000
AE017333_98	2000079	2010000
AE017333_99	2000080	2010000

AE017333-20	2000001	2110000
AE017333-21	2100002	2210000

AE017333-21 2100001 2210000
AE017333-22 2200001 2210000

AE017333-23	2300001	2310000
AE017333-23	2300001	2410000

AE017333-24

AE017333_25	2500001	2610000
-------------	---------	---------

AE017333_26	2600001	2710000
-------------	---------	---------

AE017333 27	2700001	2810000
-------------	---------	---------

AE01/333_28	2800001	2910000
AE01/333_28	2800001	2910000
AE01/333_28	2800001	2910000

AE017333-29	2900001	3010000
AE017333-30	3000001	3110000

AE017333-30	3000001	3110000
AE017333-31	3100001	3310000

AE017333-32	3200001	3310000
AE017333-31	3100001	3210000

AE017333 33

AE017333_34
3400001 3510000

AE017333_35	3500001	3610000
-------------	---------	---------

AE017333_36	3600001	3710000
-------------	---------	---------

AE017333 37 3700001 3810000

AE017333_38	38000001	39100000
AE017333_30	30000001	40100000

AEV17333_33 3900001 4010000

SOURCE
ORGANISM Pasteurella multocida subsp. multocida str. Pm70
Pasteurella multocida subsp. multocida str. Pm70
Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
REFERENCE
AUTHORS Pasteurella
May, B.J., Zhang, Q., Li, L.L., Paustian, M.L., Whittam, T.S. and
Kapur, V.
TITLE Complete genomic sequence of Pasteurella multocida, Pm70
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (6), 3460-3465 (2001)
MEDLINE 21145866
PUBMED 11248100
REFERENCE 2 (bases 1 to 13792)
AUTHORS Zhang, Q. and Kapur, V.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-2000) Department of Veterinary Pathobiology,
University of Minnesota, 1971 Commonwealth Ave., St. Paul, MN
55108, USA

FEATURES
source
1. .13792
/organism="Pasteurella multocida subsp. multocida str.
Pm70"
/mol_type="genomic DNA"
/strain="PM70"
/sub_species="multocida"
/db_xref="taxon:272843"
87. .809
/gene="PM0116"
87. .809
/gene="PM0116"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK02200.1"
/db_xref="GI:12720329"
/translation="MTTFAYGNINDYPLDSIPRELLIEKLQPOSSNQVRKHRC
RVVHFLWQLKISQKPTALKKHIDYSQSGRPQVHVDVFNISGSDWAVILRVN
TOGDSIVGDISPQKERYPALLAYFSPQAWPTQOQKEKSFYLSWCLREAVLK
SQGVGLKLSAVQHDPLLRQIQSQHCPAGQLVSHRLPFYLAFFINOTCLDHYFTW
QRGILAEITLDAIYYKVNEKT"
994. .2253
/gene="hflk"
/note="synonym: PM0117"
994. .2253
/gene="hflk"
/codon_start=1
/transl_table=11
/product="hflk"
/protein_id="AAK02201.1"
/db_xref="GI:12720330"
/translation="MENVMSLNGSDQDPWGPQKPSQRPDNEGQPNWINDNOSGR
KONNOEQSPDIEEIEFNLLKLGSGNGKRNNSHSGSLNLGKLLPIAAVIGAIW
GVSGFYTVKEARGVVRFGELHAIYQPLAWKPTFIDRVI PVNVEQYKELRTQGSML
TODENMKVETVQVRVHDPKAYLFSVTNADSLNQATDSALRYVIGHMSDDILTTG
RSVRENTWKTLNTIIEPNMGLEVVDYQSPARPEBVKDAFDIAKAEDEQRYIR
BAEYAREREPARTDAQRILEEATYKDRVVLDAKGEVERFERLLPEFKAPELLRE
RLYIOTMEKVMPTKPMVLDGNSGNNLTVPLEQLKQQTQKGSNTATSEVRPVFR
AETPNMPTQTSSTSTQENTROGRFN"
2253. .3140
/gene="hflc"
/note="synonym: PM0118"
2253. .3140
/gene="hflc"
/codon_start=1
/transl_table=11
/product="hflc"
/protein_id="AAK02202.1"
/db_xref="GI:12720331"
/translation="MRKFLTEPVIIVIAALYSIVISSEGTGRIMLRFKSVKVRDANK
VVVYVPGHFIPIDISIKILDARDITLDGADRFTVVEKKDLILDVSVKWRISDFGR
FYATGCGDYAQAQNLRRKYNDRILRSSTGRTIKDVISGTRGELMGAKALNTGFR
STAELGIEVDVVRKQINLPDEVSSTIYQVRERADEVAREHRSQGRKAAFIQADV
RKVTLILANVRNTAQELRGSDATAAKVFSDFASQEPQFYSFLRSLKAYESSFANSON
MMILKPSDSDFRFRWQAPTK"

gene
CDS
gene
CDS
gene
CDS
gene
CDS

complement (3197. .3739)
/gene="PM0119"
complement (3197. .3739)
/gene="PM0119"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK02203.1"
/db_xref="GI:12720332"
/translation="MAKRGKNKIEWNTWQEEEEEI IWVSKSEIKRDAEALKKLGK
LVLDTKAKLEKIPIEKKLELEAIELAQRLQKEARRRQLOVIGKLLRRIDVEPIQDALEK
LENKHQQQAMHLKLEMLRDLTIDNPEAITQFLEQHPDADRHRLNLRVAAQKEKAQN
KPPKAYREIFQYKDFMLED"
3834. .5189
/gene="pmba"
/note="synonym: PM0120"
3834. .5189
/gene="pmba"
/codon_start=1
/transl_table=11
/product="Pmba"
/protein_id="AAK02204.1"
/db_xref="GI:12720333"
/translation="MONTTEKSTALLKQOEBOTLREAVRYAVDLAQKAGATAEVAVTKVS
GLSVSTRLOEIBNVEFTNDGALGISVVLGNOKGNASTSDLSPEAIKNTVEAAISIAKY
ASPDSCAGLADYDLMAFDAPDLALYHAADIDVEKATATALETEQAALSYDKRIVNSEG
SAFNSHTGVRVYCNTHGMLQSYLSKRYSLSCSVIAGVEDQLERDEYTTISRDLNQLES
ASWGEHCAKTLTSLRQPKMATQVPIFLNDVATGLTSHLTAAISGSLYRKSFL
LDHKGQVLPMFHISEQPHLIGQLASSPFSEGVRTQCREIIQOGLTQYLLTYSYG
KHLGNRSTHAGGIHNLVKNNTTGLTTLTLLROMGTGLLVTLDMSCGVNLTGVDYSRG
AAGFWVENGEIQYPAVEITIAQOLKEMLRNIVAGDDIIEHRSNIQTGSLLEHLKVSG
N"
5424. .5963
/gene="hpt"
/note="synonym: PM0121"
5424. .5963
/gene="hpt"
/codon_start=1
/transl_table=11
/product="Hpt"
/protein_id="AAK02205.1"
/db_xref="GI:12720334"
/translation="MKKHVIGILISEODVRARIOTLGLREITQYQOKAVEKIVVVGLL
RGSFEMWADLRELNLVPEIEFMTTSSYSGSGTNNHDVKISKDLGGDIKGEHVLIED
IIDTGYTLQKREILKUREPASLRICITLLDKPSRREVPEVPEVWVGFSIPDEFVVGYGI
DYAQRYNLGYIGHVLEE"
5996. .6580
/gene="gmha"
/note="synonym: PM0122"
5996. .6580
/gene="gmha"
/codon_start=1
/transl_table=11
/product="Gmha"
/protein_id="AAK02206.1"
/db_xref="GI:12720335"
/translation="MYLEQIKAELENAATVNLKFIENDENNKLIIQOALLISDSFQKG
KVLSCGGSGHCDMAHFAELTGYRENPYPAIAISDVSHLSVCVSNDFGYEVFS
RYVEAVQKGVDVLFGLSTSGNSKNILNAIKAAKEKMKVIAMTCKDGGQAGIADVEI
RVPHFGVADRIQIEIHKVILHMLLIEFEAKTE"
6756. .7487
/gene="artp"
/note="synonym: PM0123"
6756. .7487
/gene="artp"
/codon_start=1
/transl_table=11
/product="Artp"
/protein_id="AAK02207.1"
/db_xref="GI:12720336"
/translation="MAIRVKNLNFYSSQALFDINLSAEGDVTVLIGPSGACKSTL
IRTLNLLVEPASGELIIANSQFDLSETKTNPKOICRLRQDVGVFQYHLWPHLTVIE
NLIEAPKVLKLSDEQARSALBLLKRLLAQADRFPLHLSSGGQQQRVAIARALMMK


```

CDS
/gene="pvuE"
3084..3848
/gene="pvuE"
/codon_start=1
/transl_table=11
/product="ferric vibrioferrin transport ATP-binding
protein PvuE"
/protein_id="BAC16543.1"
/db_xref="GI:23307113"
/translation="MLRTQNLVAYGKQTIPDLISIPGKLTALVGNPGCKSKILL
KTLVRLNKPASGVLDYDKPLSSYSGDKRLARSLSPFQILVSEGITVRKLVYGRSP
YLSHWRGLSQDQALVEQAMQDTGVTEFADKPVESLSGGQORAWTAMVNAQDITVVM
LDPEYTLDSLHQVELMKLRKMDKGTVVVVLHDLNQACRYCDHLVLEKGLQVTVQ
GTQEVLTBEGKLASVFDLQARVNDPISGTPMCVAV"
3897..3937
terminator
ORIGIN

Alignment Scores:
Pred. No.: 1.78e-27 Length: 4044
Score: 492.00 Matches: 113
Percent Similarity: 55.11% Conservative: 65
Best Local Similarity: 34.98% Mismatches: 103
Query Match: 29.11% Indels: 42
DB: 1 Gaps: 5

US-10-724-972A-6352 (1-335) x AB082122 (1-4044)
QY 31 SerSerAsnSerSerLysGluSerSerLysAspGlyValGluIleLysHis----- 47
DB 138 AGCACTGGAAGAACGACATGAAATAATAAAACAAATTACGATCAACACTTTTCAACC 197
QY 47 ----- 47
DB 198 GTTTTGTGATGCTGTTGAGCTCAGTCTTTTATGTCATTAGCGCATTTCTCAAGCGCGC 257
QY 48 -----GluGluGlyThrThrLysValProLysHisProLysArgValValVal 63
DB 258 AGCGTGCAAGATGACCAAGGCACTTTTGAGCTAGAGCATACCTCAAGTATTTGGTT 317
QY 64 LeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAla 83
DB 318 TTGGAGTTTCTGTTGTGATGCTCTGGCGCGCGTGTCTCACCAGTAGCGTGTGCC 377
QY 84 AspAspAsnLysLysAsnArgIleLysProLeuArgAspLysIleGlyLysTyrThr 103
DB 378 GATGACAAAGATGCCATACGAGTATCTCCTGCTGTGTCGCAAGATCGAACCTTGCGCA 437
QY 104 SerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLysLysProAspLeu 123
DB 438 TCGGTGCGGATGCGCTCACAAACCGAGTTTAGAGCGGATTTGCGGTCTCAAGCTGATTG 497
QY 124 IleIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaPro 143
DB 498 ATCATTTGCGGATGCCGAAGCTCATCGCGGATTTATCAAGATTTGCGCAAGATTTGCCCA 557
QY 144 ThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThr 163
DB 558 ACTTTGCTGCTAAAAAGCGGTGGAGACATATCAAGAAATCTCAAGATTCGCAACCAAAA 617
QY 164 IleSerLysAlaLeuGlyLysGluGluGluGlyLysArgLeuGluGluHisAspLys 183
DB 618 ATCGGCGTGCAATTTGCAAGCAAGCGCAATGACCAAGATTCGAGCTGCACACACAG 677
QY 184 LysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAla 203
DB 678 ACCATGCGGAGTTCACACAGCATTTTGGC-----ACACAAGAACCAATTCAGTTTGGT 731
QY 204 ValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeu 223
DB 732 GTGTTTCCGATAAAGGATGTGCTGCATAGCGGGTTTCTTACGCTGCTGCGGTTTG 791
QY 224 SerGlnLeuGlyPheLysGluAlaLeuSerAspAspValThrLysGlyLeuSerLysTyr 243
DB 791 ----- 243

```

```

DB 792 AGTACACTAGGCATCCAAAGCCCTCTGGCTCGTCTGAG----- 830
QY 244 LeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263
DB 831 ---CAAAATGCTTATATCCCAACGAGCTTCGAACTTTTACTAAAAACCAACCTGATTGG 887
QY 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeu---LysGluLeuGlu 282
DB 888 TTTACTAGTT-----GGACTTTTACTCTCAACCAATATCTGGTATGAATGGCGC 935
QY 283 LysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAsp 302
DB 936 AAAAAATCCATTATTTAAGCTGTTAACTGCGGCAAAAAGCGACAACCTGGTTCGAAGTCTCT 995
QY 303 ArgAspLeuTrpAlaArgSerGlyLeuIleSerSerGluLeuMetAlaLysGluLeu 322
DB 996 CCGGAGCTGTGTCATTAAATCTGGCATGTTGCGCGGAGAAATATGCCCGCAACCTT 1055
QY 323 ValGluLeu 325
DB 1056 GAAGCGTTA 1064

RESULT 26
AP005089/c
LOCUS
DEFINITION Vibrio parahaemolyticus DNA, chromosome 2, complete sequence, 6/6.
ACCESSION AP005089 BA000032
VERSION AP005089.1 GI:28809850
KEYWORDS
SOURCE
ORGANISM
Vibrio parahaemolyticus
Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
Vibrionaceae; Vibrio.
REFERENCE
1 Nasu,H., Iida,T., Sugahara,T., Yamaichi,Y., Park,K.S., Yokoyama,K.,
Makino,K., Shingawa,H. and Honda,T.
A filamentous phage associated with recent pandemic Vibrio
parahaemolyticus O3:K6 strains
J. Clin. Microbiol. 38 (6), 2156-2161 (2000)
20295086
10834969
2 Makino,K., Oshima,K., Kurokawa,K., Yokoyama,K., Uda,T.,
Tagomori,K., Iijima,Y., Najima,M., Nakano,M., Yamashita,A.,
Kubota,Y., Kimura,S., Yasunaga,T., Honda,T., Shingawa,H.,
Hattori,M. and Iida,T.
Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
distinct from that of V.cholerae
Lancet 361 (9359), 743-749 (2003)
22508454
12620739
3 (bases 1 to 335862)
Hattori,M., Yamashita,A., Oshima,K. and Shiba,T.
Direct Submission
Submitted (09-APR-2002) Masahira Hattori, Kitasato Institute for
Life Sciences, Kitasato University; Kitasato 1-15-1, Sagamiara,
Kanagawa 228-8555, Japan
(E-mail:hattori@genome.ls.kitasato-u.ac.jp,
URL:http://genome.ls.kitasato-u.ac.jp/,
Tel:81-42-778-8194,
Fax:81-42-778-8193)
This work was done in collaboration with Kozo Makino, Ken Kurokawa,
Katsushi Yokoyama, Takayuki Uda, Kenichi Tagomori, Masatomo Najima,
Masayuki Nakano, Yoshino Kubota, Shigenobu Kimura, Teruo Yasunaga,
Takeshi Honda, Hideo Shinagawa, Tetsuya Iida (Osaka University),
Yoshio Iijima (Kobe Institute of Health), and supported by the
Research for the Future Program of the Japan Society for the
Promotion of Science. This clone was isolated from a patient
presenting with acute gastroenteritis.
Location/Qualifiers
1. 335862
/organism="Vibrio parahaemolyticus"
/mol_type="genomic DNA"
/strain="O3:K6"
FEATURES
source

```

```

/sub strain="RIMD 2210633"
/db_xref="taxon:670"
/chromosome="2"
/complement (55. .1971)
/gene="VPA1449"
/complement (55. .1971)
/gene="VPA1449"
/notes="similar to GB:AAF96089.1 (AE004358) percent
identity 71 in 638 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="methyl-accepting chemotaxis protein"
/protein_id="BAC62792.1"
/db_xref="GI:28809851"
/translation="MLKLHSLSIKQKVLGITFAVLASTIIVGMAQRHARDVLSHRL
IDLEPAMLQAIINTEIDREVMQOAKOLATNEFVVEALKNTHDPQSFSTOLVQQLN
NIKSOGLNDQAVANRKYAYVMQNGFLRNLNSQDAWPGFTSSGRETSVSVFQAN
GEVKMTNYQDLNGISNGLSKSDMDVSLNSFQIEDTGIFLTNKEGDIQIHRQOG
KNKTSIAQLGNSVANQLLNKNFNLNVEFGKDNFIASLYPSNMFWIGVVPDEV
FADLNATGQOMITIVVALVFLMGVLLANSITKPIRLIADFTDLGQCEGLDSQRI
EIRGNDEIAQLSGFNGFTFKIHAATKVEVSLTSGALSQAADSVSSEKATSYDMSQOR
DQIIVAAINQMGATISIAANAATAADTANQASNTOTGREVVMAKEVLSRLADD
EVNTNIWVQLASTTIGSILGIVRDISETQNLALNALEAARAGEQGRGPAVVAD
EVRNLASRTADSFEEIORMINQLOSDAQDAVSMEAGKAVTFEGVASTDEAVEVLNI
SERLIDSRNTQVATATEQSTVHTINQNIIEINAINEMTTATAEELAGASRLQCE
LSRLDMVGSFKL"
/complement (2136. .2288)
/gene="VPA1451"
/complement (2136. .2288)
/gene="VPA1451"
/notes="unknown"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC62793.1"
/db_xref="GI:28809852"
/translation="MLWTLDIHLIFVLSIGHCQSGKSCSLSKFISILGQRKGLQAK
KNAESA"
2287. .3243
/gene="VPA1450"
2287. .3243
/gene="VPA1450"
/notes="similar to GB:AAF96088.1 (AE004357) percent
identity 79 in 318 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="MoxR-related protein"
/protein_id="BAC62794.1"
/db_xref="GI:28809853"
/translation="MHKSNFETLQHYLESQIIIGHDLVKQLLIALLADGHILVGGPPG
LAKTRAVKSLSDVCVEGFHRIQPTFDLLPADLTGTDIFRPETGEFTFQSGPIFNSLIL
ADEINRAKQVQAMLEMAEKQVTAGRHATLPELFLVMATQNPIDQEGTYPILPEAQ
LDRELLHLEVDPDAHELAELIRINRGEAKGECIERPSLSQQDIPTARQEVLDIHW
EAIQYIVRLVMATRASSEYSELDKWLAMGVSPRATIALDRCARAHAWLAGRDFVSP
EDVQAMYPVLRHLLLSYHAQAEQVTSNQVIDKLLNLVGS"
3258. .4184
/gene="VPA1452"
3258. .4184
/gene="VPA1452"
/notes="similar to GB:AAF96087.1 (AE004357) percent
identity 50 in 309 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC62795.1"
/db_xref="GI:28809854"
/translation="WSANHPLPHANGVTLTLELLQYRTQSVNWLPPAQSIWSQMSG
NHPISROKGRGDMEVRYQAGDDIRSIDMRVTARTKAHTKFAEDKEQAVILYLDL
CSSLHFGSKTVLSVQLAHFASVLIWTLAKKORIGAVIDDGHQCLEFRPSSLQKQGL

```

```

RILNAIVDTHNAQLTQDPNHETRMPTYSOVLETLHTLTPKGSSELILLSDFAHVBEREL
TOLRRLLKQHNVSRAIQFYDPLERGETDFRQAKASDGRRSQWNFNGTGORKKALETHF
TQHQESIKQOCHSMAIFNTLSSGWPLIQOLS"
4225. .4701
/gene="VPA1453"
4225. .4701
/gene="VPA1453"
/notes="similar to GB:AAF96086.1 (AE004357) percent
identity 51 in 151 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC62796.1"
/db_xref="GI:28809855"
/translation="MTTNQSNLEPLIILPNAPDMFPLAWGWMALAAI VLFVLGVYL
YAKWRKRLALKTKLTLTQSVTPHTPSSALEILQAAALSYPREDIAPLTGMAWE
FIDSHTKTRFVMDKQOOWAALYQKSPQHQHDLIDDCAFWITHALPPKKKANTRD"
4712. .5698
/gene="VPA1454"
4712. .5698
/gene="VPA1454"
/notes="similar to GB:AAF96085.1 (AE004357) percent
identity 65 in 317 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC62797.1"
/db_xref="GI:28809856"
/translation="MTDLSQMLNIEFVMWALFLLPLLIYKLLPEETRQABIKLAY
LPESKNSRKPQWLQKSLAYGVWTLIVACAPVWFGEVPEPKYRDLMLVVDLSGS
WQKEMNLIDGEYIDRLSAVKKVLSDFAVRKGRDLGVLFGDHAYLQTLTADRTQVI
QIKQTVIGLVGQRTAIGDIGLGTFTFVSDAPQRMVLLSGNTAGVLDPIEAAE
IAKYNATITVGVGAGEMVMDFFMTRKVDAAADLEQTLTKIAEMTSGQYFRADA
FOLEKIYDITINKLEPVSDTQTRPQSEWFPYPLSAALVLSMLLLFLLRRKHG"
5691. .7547
/gene="VPA1455"
5691. .7547
/gene="VPA1455"
/notes="similar to GB:AAF96084.1 (AE004357) percent
identity 43 in 653 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC62798.1"
/db_xref="GI:28809857"
/translation="MASEFLNPNYLLGLITIPFVLLINQKFRAKKSALIAPIHLAKML
GQTTQSKPYFIWGLANTIACALGPSQSNTRPSELSQNRMLVLDMSRSMYASDI
KPNLAQTRYKALDLLPKWKEGATGLIYAGDAYLSPLTTDASTLAGI IENLSPELM
PFQSNLPAIAIELSQFSQAGANQGDIVLADDLDSSELSLDLVKGNIRVSVLA
IGTAGAPIALPDGSLKTTQGSTVWAKTHLKNLTQTLANKTGGMFVPIQHNNRVDISI
AAYTNTGNLAAKQNOEVKTDLSRLNGFWLLPLLLPTVLPFRGVIVWFAALLPM
TWTDAQANPLNADQAAANLYKQGEYQAKNLTDFDSWQGAASYQTGDYBAALKAFS
NDPSQTRYNLANAQNGLEADKQTKLLEENPDFEAAKNLSVVEEKLKQOQQQ
QOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQ
QOQQOQNRTKQEQQQOQQOQDAENKEKESEHKAQKQKAEVEKQGEQKQGPMPQ
AQAPQEXSHODDPFRLEAVESARDPSFLIRAQMLQAOQKQRPQSQKSW"
7555. .9198
/gene="VPA1456"
7555. .9198
/gene="VPA1456"
/notes="similar to GB:AAF96082.1 (AE004357) percent
identity 44 in 200 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC62799.1"
/db_xref="GI:28809858"
/translation="MKKGVLVFLIVASLLSSFSVMAQSLQASVNKTEVAKNEVINLR
IMADSEVASDAIDFSVLEKDFLQGPRIYKSSNNNGRKRYORTWESIAPMKREGIIT

```

IPSPADGKHTPIQLKVTANKSAPDLDDLFSPNMSVDNHTLYPOOSANLRMLIILKA
DTRLDNPQVPRPIEGMKLEVPVEMQOQQRVIAGLEVTVVEQSRLTAEOQGTFTLL
GFLQTSYIYSGSLSTKIMPISTKVEQMPITVKAIPSEFGKSWLPGSALQNTQSQW
DQGNLTSAVNTVNNKQSSITRTIQKARGTQAEYLPRTIMDYFNLSLRYVEQOQFD
TARDGTIVNTVQVLIPTEAGEFTLPGYTLNWMDSKSDKAQANLSLXINVEQSDAG

Alignment Scores:
Pred. No.: 7,81e-25 Length: 335862
Score: 487.50 Matches: 110
Percent Similarity: 58.52% Conservative: 72
Best Local Similarity: 35.37% Mismatches: 106
Query Match: 28.85% Indels: 23
DB: 1 Gaps: 5

US-10-724-972A-6352 (1-335) x AP005089 (1-335862)

QY	16	LeuLeupheValLeuileAlaThrAlaAaCysglyAsnAenSerSerSerAenSerSer 35
DB	229956	GTTTGCTGATGCTAGCTAGCTTGTATGTCATTAGCGCATTTCTCAGCGCGC 229897
QY	36	LysGluSerSerLysAspGlyValGluileLysHisGluGluGlyThrLysValPro 55
DB	229896	-----AGCGTGCAAGATGAGCAAGGCACTTTTGAGCTAGAA 229861
QY	56	LysHisProLysArgValValValLeuGluTyrrSerPheValAspAlaLeuValAlaLeu 75
DB	229860	CGCATCTCAACGTAATTTGGAGTTTGGTGTGTCGCGCTGCGCGCGCTC 229801
QY	76	AspValLysProValGlyLeAlaAspAsnLysLysAsnArgilleLysProLeu 95
DB	229800	GATGCTCACCGGTAGCGGTGGCGATGACCAACGATGCCATCAGCATCTCCTGCTGTT 229741
QY	96	ArgAspLysIleGlyLysThrThrSerValGlyThrArgLysGlnProAsnLeuGluGlu 115
DB	229740	CGTGCAAGATCGAACCTTGGCAATCGCTCGCATCGCTCACAACCGAGTTTAGAGCGC 229681
QY	116	IleSerLysLeuLysProAspLeuileleAlaAspAsnAenArgHisLysGlyLetyr 135
DB	229680	ATTGCGGTGCTCAAGCCTGATTTGATCATTTGCCGATGCGCAACGTCATTTGCGCGGATTTAT 229621
QY	136	LysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrrAen 155
DB	229620	CAAGATTTGCGAGCAATTTGCGCAACTTTGCTGTAAAGCCGCTGGTGAAGCATATCAA 229561
QY	156	GluAenIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGluGlyLys 175
DB	229560	GAATAATCGAATCCGCACAAAATAATCGCGTGGCAATGGCAAGCGCAAAATGAGC 229501
QY	176	LysArgLeuGluGluHisAspLysLysIleGluGluTyrrLysLysGluIleThrMetAsp 195
DB	229500	CAACGAATCGAGTTGCACAAACAGACCATGGCGGAGTTCAAAACAGCATTTTCGC----- 229447
QY	196	LysAsnGlnLysValLeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHisProSer 215
DB	229446	ACACAAGAAACCAATTCAGTTTGGTGTGTTCCGCAAGGAGATGGTTGTGATAGCCCG 229387
QY	216	AsnSerTyrrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAsp 235
DB	229386	GTTCCTTACGCTGGCGGTTTGACACTCTAGGATCCAAAGCCCGTGGCTCCATCT 229327
QY	236	ValThrLysGlyLeuSerLysTyrrLysGlyProTyrrLeuGlnMetAenThrGluThr 255
DB	229326	GAG-----CGCAATGCGTATATCCCAACCAAGCTTTGACACTT 229291
QY	256	LeuSerGlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSerAsnGlu 275
DB	229290	TTACTAAAAACCAACCTGATGCTTACTAGTT-----GGACTTTACTCTCAA 229243
QY	276	ProSerLeu---LysGluLeuGluLysAspProValTrpLysLysLeuAsnAlaValLys 294
DB	229242	CCAAATATCGTGAATGCGGTAGAAATCCATTTATTTAGTTGCTTAAGCTGCGTCAAA 229183
QY	295	AsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArgGlyLeuIleSer 314

DB 229182 AAAAAACAACCTGGTTCGAAGTCTCTCCGGAACCTGTGTCATTAAATCGCGCATGTTAGCG 229123
QY 315 SerGluGluMetAlaLysGluLeuValGluLeu 325
DB 229122 GCGGAAGAAATTCGCCGCAACCTTGAAGCGTTA 229090

RESULT 27
BX950851_11/c
WPCOMMENT
Sequence split into 51 fragments LOCUS BX950851 Accession BX950851

Fragment Name	Begin	End
BX950851_00	1	110000
BX950851_01	100001	210000
BX950851_02	200001	310000
BX950851_03	300001	410000
BX950851_04	400001	510000
BX950851_05	500001	610000
BX950851_06	600001	710000
BX950851_07	700001	810000
BX950851_08	800001	910000
BX950851_09	900001	1010000
BX950851_10	1000001	1110000
BX950851_11	1100001	1210000
BX950851_12	1200001	1310000
BX950851_13	1300001	1410000
BX950851_14	1400001	1510000
BX950851_15	1500001	1610000
BX950851_16	1600001	1710000
BX950851_17	1700001	1810000
BX950851_18	1800001	1910000
BX950851_19	1900001	2010000
BX950851_20	2000001	2110000
BX950851_21	2100001	2210000
BX950851_22	2200001	2310000
BX950851_23	2300001	2410000
BX950851_24	2400001	2510000
BX950851_25	2500001	2610000
BX950851_26	2600001	2710000
BX950851_27	2700001	2810000
BX950851_28	2800001	2910000
BX950851_29	2900001	3010000
BX950851_30	3000001	3110000
BX950851_31	3100001	3210000
BX950851_32	3200001	3310000
BX950851_33	3300001	3410000
BX950851_34	3400001	3510000
BX950851_35	3500001	3610000
BX950851_36	3600001	3710000
BX950851_37	3700001	3810000
BX950851_38	3800001	3910000
BX950851_39	3900001	4010000
BX950851_40	4000001	4110000
BX950851_41	4100001	4210000
BX950851_42	4200001	4310000
BX950851_43	4300001	4410000
BX950851_44	4400001	4510000
BX950851_45	4500001	4610000
BX950851_46	4600001	4710000
BX950851_47	4700001	4810000
BX950851_48	4800001	4910000
BX950851_49	4900001	5010000
BX950851_50	5000001	5064019

Continuation (12 of 51) of BX950851 from base 1100001 (BX950851 Erwinia carotovora subsp
Alignment Scores:
Pred. No.: 1.16e-24 Length: 110000
Score: 477.50 Matches: 110
Percent Similarity: 54.01% Conservative: 65
Best Local Similarity: 33.95% Mismatches: 120
Query Match: 28.25% Indels: 29
DB: 1 Gaps: 6

US-10-724-972A-6352 (1-335) x BX950851 11 (1-110000)

QY	8	LeuLysIleLeuSerValIleGlyLeuPheValLeuIleAlaThrAlaAlys---	26
Db	105548	ATTGAGGTATTCCTTCTCTACTTCGGCTTGCAATGATTACAAATGCTTTGTCGTG	105489
QY	27	-----GlyAsnAenSerSerAenSerSerLysGluSerSerLysAspGlyVal	43
Db	105488	TTCCGACTGGGTGCG-----GCCAATGCAAGTA	105462
QY	44	GluIleLysHiGluGluGlyThrThrLysValProLysHisProLysArgValValVal	63
Db	105461	ACGGTGCAGGATGAACAGCGTTCCTTACGCTCGACACGCCCTCAGCGCATTTGGTG	105402
QY	64	LeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAla	83
Db	105401	CTGGAACTCTCTTTTGGCCGATGACTGGCCGCGATAGACGTGAGCCCATCGCATCGCC	105342
QY	84	AspAspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThr	103
Db	105341	GATGACACGACTCCAGACCGGATCTCGGCAGAGTCCGTGAACACCTAAACCGTGGCAC	105282
QY	104	SerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeu	123
Db	105281	TCAGTGGGAACCGGTGCGCAGCGGACCTCGAAGCCATCAGCGCTGAAGCCTGATTTG	105222
QY	124	IleIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaPro	143
Db	105221	ATTATCGCGACAGCAGTGCACACCGGTATTTATACGGCACTAAGGCTATCGTCCG	105162
QY	144	ThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThr	163
Db	105161	GTACTGTCTGCTGAATATCAGCAACCACTATGATGAACACCTGCTTCGCGGAATC	105102
QY	164	IleSerLysAlaLeuGlyLysGluGluGlyLysLysAspLeuGluGluHisAspLys	183
Db	105101	ATCGCAAGGTATTAGGGAAGACACAGTCGATGCGAGGCTCGTCTTGCTAAGCATCGCGC	105042
QY	184	LysIleGluGluTyrLysLysLulleThrMetAspLysAsnGlnLysValLeuProAla	203
Db	105041	ACCATGAAGGCTACGACACGAG-----CTTCCAAAGCACGAGCGTCGCTTCGCG	104988
QY	204	ValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeu	223
Db	104987	ACCTCGCGCGAAGCAGCAGTTTAACTGTATTCCAGCGAGGCTACACCGGACGCTGTG	104928
QY	224	SerGlnLeuGlyPheLysGluAlaLeuSerAspAspValThrLysGlyLeuSerLysTyr	243
Db	104927	GCTGCTCTGGGTCTG-----AGCGTGGCGAAACCATCAAT-----	104892
QY	244	LeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg	263
Db	104891	--AATGCCCGATGGCTCCATCAATTTGGACAGCTGCTGGCGATTATTCACAGTGG	104835
QY	264	MetPheIleMetThrAsnLysAlaSerAsnGluProSerLeuLysGluLeuLys	283
Db	104834	CTGATTGTGACGCACTACTCG-----GAAGAAAGTATTGTGAACCGCTGGCAGCAG	104784
QY	284	AspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg	303
Db	104783	GATACGCTCTGGAACATGCTGGGAAGCACAGCAAAACAGCAGATTGCGCGCTGCAGCT	104724
QY	304	AspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuVal	323
Db	104723	AACGCTGGCGCGGATGCGCGCATCTTTGACGGGAACCGCTCGGCACCGATACGGTG	104664
QY	324	GluLeuSerLys	327
Db	104663	AAGATTTTTTAA	104652

RESULT 28
BX950851 12/c
WPCOMMENT

Sequence split into 51 fragments				LOCUS	Accession
Fragment Name	Begin	End			
BX950851_00	1	110000			BX950851
BX950851_01	100001	210000			
BX950851_02	200001	310000			
BX950851_03	300001	410000			
BX950851_04	400001	510000			
BX950851_05	500001	610000			
BX950851_06	600001	710000			
BX950851_07	700001	810000			
BX950851_08	800001	910000			
BX950851_09	900001	1010000			
BX950851_10	1000001	1110000			
BX950851_11	1100001	1210000			
BX950851_12	1200001	1310000			
BX950851_13	1300001	1410000			
BX950851_14	1400001	1510000			
BX950851_15	1500001	1610000			
BX950851_16	1600001	1710000			
BX950851_17	1700001	1810000			
BX950851_18	1800001	1910000			
BX950851_19	1900001	2010000			
BX950851_20	2000001	2110000			
BX950851_21	2100001	2210000			
BX950851_22	2200001	2310000			
BX950851_23	2300001	2410000			
BX950851_24	2400001	2510000			
BX950851_25	2500001	2610000			
BX950851_26	2600001	2710000			
BX950851_27	2700001	2810000			
BX950851_28	2800001	2910000			
BX950851_29	2900001	3010000			
BX950851_30	3000001	3110000			
BX950851_31	3100001	3210000			
BX950851_32	3200001	3310000			
BX950851_33	3300001	3410000			
BX950851_34	3400001	3510000			
BX950851_35	3500001	3610000			
BX950851_36	3600001	3710000			
BX950851_37	3700001	3810000			
BX950851_38	3800001	3910000			
BX950851_39	3900001	4010000			
BX950851_40	4000001	4110000			
BX950851_41	4100001	4210000			
BX950851_42	4200001	4310000			
BX950851_43	4300001	4410000			
BX950851_44	4400001	4510000			
BX950851_45	4500001	4610000			
BX950851_46	4600001	4710000			
BX950851_47	4700001	4810000			
BX950851_48	4800001	4910000			
BX950851_49	4900001	5010000			
BX950851_50	5000001	5064019			

Alignment Scores:

Pred. No.:	1.16e-24	Length:	110000
Score:	477.50	Matches:	110
Percent Similarity:	54.01%	Conservative:	65
Best Local Similarity:	33.95%	Mismatches:	120
Query Match:	28.25%	Indels:	29
DB:	1	Gaps:	6

US-10-724-972A-6352 (1-335) x BX950851 12 (1-110000)

Qy	8	LeuLyarIleuSerValIleGlyLeuLeuPheValIleuLeaThrAlaAaCys---	26
Db	5548	ATTAGGAGTTACCGTATGTTTGTCTACTTCGGGTGCTACATTAACAATGCTTTTGTCG	5489
Qy	27	-----GlyAsnAenSerSerAsnSerSerLySgluSerSerLySAspGlyVal	43
Db	5488	TTCGGACTGGTGGCGC-----GCCAATGCAGTA	5462

Db 7829 CAGGATACCTGTGGCAGATGATCAGACGACAGCAGCATAGATGACGCGGTAGAC 7770
Qy 303 ArgAspLeuTrrAlaAaGserArgGlyLeuIleSerSerGluGluMetAlaLysGluLeu 322
Db 7769 AGCAACACCTGGGCTGTGTCGGGGATATTTCGCCGTGAACCATGGCAGCGATGCG 7710
Qy 323 ValGluLeu 325
Db 7709 GTAAAAATC 7701

RESULT 30
BX571873
LOCUS
DEFINITION
Photorhabdus luminescens subsp. laumondii T701 complete genome;
segment 15/17.

ACCESSION
BX571873 BX470251
VERSION
BX571873.1 GI:36787441
KEYWORDS
complete genome.

SOURCE
Photorhabdus luminescens subsp. laumondii T701
Photorhabdus luminescens subsp. laumondii T701
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Photorhabdus.

REFERENCE
AUTHORS
Duchaud, E., Rusniok, C., Frangeul, L., Buchrieser, C., Taourit, S.,
Bocs, S., Boursaux-Eude, C., Chandler, M., Dassa, E., Derose, R.,
Derzelle, S., Freyssinet, G., Gaudriault, S., Givaudan, A., Glaser, P.,
Medigue, C., Lanois, A., Powell, K., Sigulier, P., Wingate, V.,
Zouine, M., Boemare, N., Danchin, A. and Kunst, F.
Complete genome sequence of the entomopathogenic bacterium
Photorhabdus luminescens
Nat. Biotechnol. 11 (1) (2003) In press

JOURNAL
REFERENCE
AUTHORS
Duchaud, E., Frangeul, L., Rusniok, C. and Kunst, F.
Direct Submission
Submitted (23-APR-2003) L. Frangeul, Institut Pasteur, Genopole, 25
rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE. E-mail:
lfrangeu@pasteur.fr, fkunst@pasteur.fr
LOCATION/Qualifiers
1. 349107

FEATURES
source
/organism="Photorhabdus luminescens subsp. laumondii T701"
/mol_type="genomic DNA"
/strain="T701"
/db_xref="taxon:243265"
complement (80. .3225)
/gene="tccc1"
/locus_tag="plu4167"
complement (80. .3211)
/gene="tccc1"
/locus_tag="plu4167"
/function="Unknown, probable insecticidal toxin"
/note="Highly similar to insecticidal toxin complex
protein TccC of W14."
/codon_start=1
/transl_table=11
/product="Insecticidal toxin complex protein Tccc1"
/protein_id="CAE16539.1"
/db_xref="GI:36787442"
/translation="MSTPDTALYTQPTVSVLDNRGLSIRDIGFHRVVGIDITRVT
RHOVDARGYLNSIDPLRYDAKQDQNSVKNFVQYDLGHALTRESVDAGRTVALND
IEGRPVMTNATGVRQTRRYEGNTLPGRLLSVSEQVDFQETSQVTERFIAENTAEK
EYNLSGCIRHYDAGTVRLMSQSLAGAILSQSHQWLAEGQEAQNSGDDETVMQMLA
SEVYVSTSTNIGAILTQDAKNIQRLAYDIAGQLKSWLTVKGQEOVIVKLSLW
SAAGKLEEHGCVITEYSYEPETQRLMGTTRRAESQSEARVLODLRYEVDVGN
VISVHNDAAEAFWRNOKVEPENRYVYDLSYQLMSATGRENANIGQSQNSLPTTDP
PADNTVNTQRTYSYRGGLNQLVQIRHSSPATQNSYTEMVTSRSNRNLVSLTITDP
TVQDGLFDAGHQKTLIPQNLDMRGEQLQRTVPVSRNSDSEWRYSSDGVRLIK
VSEQTGNSQVQRTVLPGLRLITGADTKTQNLQVITVGEAGRAQVRVLHWSGK
PADIDNNQRYSYDNLGSSQIELDSEQLDSQEEYPIGGTAVMAAKNOTEASVYFI
PSKGRDADLYGYGYRYPQWRLSDLPAGTVGLNVRNRPILTDHGL
APSNRNNTFWFASFLFRKPDGMSMRGOKIGRALAGGLAIGGLAAIAGTAGA
ALPVLGAAVAGGIGALMGVNSLLEKGGALLARLVQGRKSLVQSGAAGAGASGA
AAYCARAQGVGASAGAVTCAVGSVINNADRGIGGAGAGAVGTITDMLGTSSTLT
HEVAAAAGGAAGGMITGTGSTRAGIHAGITGYTGSWIGFGLDVASNPAGHLANYAG

gene

CDS

RBS

gene

CDS

YAAAGLGAEMAINRVIGGGFLGRLLGRAVSPYAAGLARQLVHFSTIARVPFPIPSVLGG
LAGGIGTGLHRVWGRSWSRGLSAAAGSDHLAGMIGNQIRGRVLTTGTGIANADYG
TSAIGARRVISL"
complement (3220. .3225)
/gene="tccc1"
/locus_tag="plu4167"
complement (3345. .8051)
/gene="tccc1"
/locus_tag="plu4168"
complement (3345. .8039)
/gene="tccc1"
/locus_tag="plu4168"
/function="Unknown, probable insecticidal toxin"
/note="Highly similar to insecticidal toxin complex
protein tccB of W14 Putative transmembrane protein."
/codon_start=1
/transl_table=11
/product="Insecticidal toxin complex protein TccB1"
/protein_id="CAE16540.1"
/db_xref="GI:36787443"
/translation="MLSTMEKQLNESQDALVTGYMNFVASTLKGVDQPVTVEDLYE
YLIDPEVADEVETSRVQAIAISIQYMTLVNGSEPGRAQAMELSTANEWRDNDQV
VNSLVVSGYINQDKFDOAIYVFIGRTTKPYRYWRNDLSDRVDODVGLYNERFEA
NDWQETPLSGDVTLEHTVRPVFVNDRLYVANVERDPAQVQADGKNIGKTHAYSIK
FGYKRYDDTPATNTTLMTQADSESSTQRSLIDESKTLTKRKNLLATDTSIDP
TEETDSNPGRLMLGVFNKFTGTPHQNPAPVGYLYCDSAFNHHLRPLSDFLST
YRDETNDQCQFAYVKKYVITIKSGPTEDPENTGWGVDDLKLGTTGAYVYID
QDGLTHQITTTNGFVNRHIVYNDIVVDSKSGYFTWSGNEGYLDVHDGNYTFH
NAAINYPAGYGGSGVNGTWALEQVNEGWAISPLDDAHTTITVQGSYIAMEGETPA
SYLPIPDGTMLLDMSDKINPAISLKLKLESVTSFDPWPLTTIKPFSKIADNRKFKQE
INAEADGRNLFKHYSEQTFQASGATPTTTLSEADFSTDPDKNVLQVCLNVDWH
YDRSGRGASWSKNWVYVDELQSKAPDAIPRLISRYDSKRGVLQVDFWTSLLP
AKRNTLTVFTLIEKANGLDLSLDYTLQADPSLEADLVTDCKSEPMDFDGSNGLYF
WELFPHLPFLVATRFANQSPQAOKSLHYIPDPAMKNKPNAPVWVRPLVEVHSD
LSRHLCGDPIDPTQAYAPVYKAVFIAYVSNLJAQGDWYRQLTROGLTOARVYIN
LAELGPRPDVLSSTWTRQTLAAGQKLTALRDFEHQLANSVTLPALFENNLSH
LKLADNGYFNEPLVNLMSHMDTLARLNLNHLNLTVDGKPLSLPLYPADPVALLA
QRAQGTUANGSVGAILTVPPYRFSAPVAVTGTSTLTSFQNLISLLESERAQOE
ELAQQLDMSYAITLQOALDGLADRLALLAQATAQQRHDHYTLTYONNISAE
QLVMDTQTSQSLISSSTQVOTAGALKVIPNIFGLDGSRYEGVTEAIAIGLWAG
QATSVAERLATTENRRRREMOIQOYQAOSEVDALOKOLDALAVREKAAQTSIQQA
KAQVQIRITMLTYLTTRFTQATLQWLSQSLAYYQAYVSLCSAQACQWGLGD
DYATFTQGTWHDYRGLQVGETLQNLNQEAAYLVRRHRRNLVTSLSKSLGD
DGFGLKTEGKVDPLSEKLFNDNDYPGHYRQIKTVSVTLFTLIGPQNVKATLTQTS
SSTLLTADINVKRLNDPTGEGDATHIVTNLRASQVQVALSSGINDAGSFELRLEDER
YLSFEGTGAISKWTLTFPRSVDEHIDDKTKADEMQAALLANMDDVLQVHYTACDGG
ASFANQVKTLF"

RBS

gene

CDS

complement (8046. .8051)
/gene="tccc1"
/locus_tag="plu4168"
complement (8134. .11034)
/gene="tccc1"
/locus_tag="plu4169"
complement (8134. .11034)
/gene="tccc1"
/locus_tag="plu4169"
/function="Unknown, probable insecticidal toxin"
/note="Highly similar to insecticidal toxin complex
protein TccA of W14."
/codon_start=1
/transl_table=11
/product="Insecticidal toxin complex protein TccA1"
/protein_id="CAE16541.1"
/db_xref="GI:36787444"

/translation="MNQLASPLISRTBEIHNLPKLTDLGYTSVDFVVRMPREPIRE
HRADLGRSAGKMYDLAGVYAHVHLHFRNLSSEAVQFSRSPESVSGPDYANQFLDA
NTWKDKAPSGPEANDPAVYALTHYQLADQKSGASAIMNTLVERRPDIGALLIN
DAKNEVYIPLQVLNLTLSKAIQKSLDLEAVNAKLSTTRYPNNLPYHYGQIOIOT
AQVSLGTLLQDITLPTQTLDPQNFQWATAKGLSDATAGALTRLQIMASOLSPEOQIM
VETVQGNFYQGLTANFSFNDLTDRSLTQVVELMCLSTVSGSVYVKSND
VSSGDTTATFAYGARFIHAGKQATLSRSRGAHFPALVNLMDKDLORINTKRL
QKWLNPYEDIDLLVTSAMDARETGNLTALMNDNTLRLMLGVFKHYQAKYGTAKQFAGH

[illegible]

Alignment Scores:

Alignment Scores:	5.99e-24	Length:	349107
Pred. No.:	476.00	Matches:	102
Score:		Conservative:	60
Percent Similarity:	57.24%	Mismatches:	107
Best Local Similarity:	36.04%	Indels:	14
Query Match:	28.17%		

AR354237	AR354237	Sequence 355 from patent US 6593114.	668 bp	DNA	linear	PAT 17-AUG-2003
LOCUS						
DEFINITION						
ACCESSION	AR354237					
VERSION	AR354237.1	GI:33760321				
KEYWORDS						
SOURCE		Unknown.				
ORGANISM		Unknown.				

Unclassified.
1 (bases 1 to 668)
REFERENCE Kunsch,C.A., Choi,G.H., Barash,S., Dillon,P.J., Fannon,M.R. and
AUTHORS Rosen,C.A.
TITLE Staphylococcus aureus polynucleotides and sequences
JOURNAL Patent: US 6593114-A 355 15-JUL-2003;
FEATURES Location/Qualifiers
source 1. .668
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 3.9e-27 Length: 668
Score: 475.00 Matches: 91
Percent Similarity: 86.89% Conservative: 15
Best Local Similarity: 74.59% Mismatches: 16
Query Match: 28.11% Indels: 0
DB: 6 Gaps: 0

US-10-724-972A-6352 (1-335) x AR354237 (1-668)

QY 210 LeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLys 229
Db 2 TTATTAGCACATCCAACTATTCTGCTGAAATTCAGAGAGATTGGAACAATTTTAAACGAATAGGATTTAAA 61
QY 230 GluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeu 249
Db 62 AATGCATTAAAGTACGATGTAACAAGGTTTAAGTAAATTTTGAAGACCTTACTTA 121
QY 250 GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheileMetThrAsn 269
Db 122 CAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGCGTATGATCATTTATGACAGAT 181
QY 270 LysAlaSerSerAsnGluProSerLeuLysGluLeuLysGluLysAspProValTrpLysLys 289
Db 182 CATGCTAAAAAGATTCTGCTGAAATTCAGAGAGATTGGAACAATTTTAAACGAATAGGATTTAAA 241
QY 290 LeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSer 309
Db 242 TTGAATGCAGTTAAAAATTAATCGCGTGATATGTTGACCGGTATTTGGGCAAGATCT 301
QY 310 ArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAsp 329
Db 302 CGTGCTTAAATTTCTTCTGAAGAATGCTTAAAGAATTTGTTGAATTTATCAAAAAAGAA 361
QY 330 SerLys 331
Db 362 CAAAAG 367

RESULT 32
LOCUS AR535793 668 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 355 from patent US 6737248.
ACCESSION AR535793
VERSION AR535793.1 GI:53927010
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 668)
AUTHORS Kunsch,C.A., Choi,G.A., Barash,S.C., Dillon,P.J., Fannon,M.R. and
Rosen,C.A.
TITLE Staphylococcus aureus polynucleotides and sequences
JOURNAL Patent: US 6737248-A 355 18-MAY-2004;
FEATURES Location/Qualifiers
source 1. .668
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 3.9e-27 Length: 668

Score: 475.00 Matches: 91
Percent Similarity: 86.89% Conservative: 15
Best Local Similarity: 74.59% Mismatches: 16
Query Match: 28.11% Indels: 0
DB: 6 Gaps: 0

US-10-724-972A-6352 (1-335) x AR535793 (1-668)

QY 210 LeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLys 229
Db 2 TTATTAGCACATCCAACTATTCTGCTGAAATTCAGAGAGATTGGAACAATTTTAAACGAATAGGATTTAAA 61
QY 230 GluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeu 249
Db 62 AATGCATTAAAGTACGATGTAACAAGGTTTAAGTAAATTTTGAAGACCTTACTTA 121
QY 250 GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheileMetThrAsn 269
Db 122 CAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGCGTATGATCATTTATGACAGAT 181
QY 270 LysAlaSerSerAsnGluProSerLeuLysGluLeuLysGluLysAspProValTrpLysLys 289
Db 182 CATGCTAAAAAGATTCTGCTGAAATTCAGAGAGATTGGAACAATTTTAAACGAATAGGATTTAAA 241
QY 290 LeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSer 309
Db 242 TTGAATGCAGTTAAAAATTAATCGCGTGATATGTTGACCGGTATTTGGGCAAGATCT 301
QY 310 ArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAsp 329
Db 302 CGTGCTTAAATTTCTTCTGAAGAATGCTTAAAGAATTTGTTGAATTTATCAAAAAAGAA 361
QY 330 SerLys 331
Db 362 CAAAAG 367

RESULT 33
LOCUS ECOFECBCDE 4842 bp DNA linear BCT 26-APR-1993
DEFINITION K-12 fecA gene, 3'end; fecB,C,D, and E genes, complete cds's.
ACCESSION M26397
VERSION M26397.1 GI:145923
KEYWORDS
SOURCE Escherichia coli
ORGANISM Escherichia coli

REFERENCE 1 (bases 1 to 4842)
AUTHORS Staudenmaier,H., Van Hove,B., Yaraghi,Z. and Braun,V.
TITLE Nucleotide sequences of the fecBCDE genes and locations of the
proteins suggest a periplasmic-binding-protein-dependent transport
mechanism for iron(III) dicitrate in Escherichia coli
JOURNAL J. Bacteriol. 171 (5), 2626-2633 (1989)
MEDLINE 89213950
PUBMED 2651410

COMMENT Original source text: Escherichia coli (strain K-12) DNA.
FEATURES Location/Qualifiers
source 1. .4842
/organism="Escherichia coli"
/mol_type="genomic DNA"
/strain="K-12"
/db_xref="taxon:562"
1. .136
/gene="fecA"
1. .136
/gene="fecA"
/codon_start=2
/transl_table=11
/protein_id="AAA23761.1"
/db_xref="GI:145924"
/translation="DLNLAFGVKNIFDQYFIRSYDDNNKGIYAGQRTLYMQSLKF"

gene

CDS

gene

CDS	/gene="fecB"		85	AspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSer	104
	181..1083		370	GATAACGATGCAAAACGATCTCGCCGAGTGGCGACCTGAAACCGTGGCGATCC	429
	/codon_start=1		105	ValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIle	124
	/transl_table=11		430	GTCGGAACCGCGCGAGCCCTGGAGCCATTCGCGCTCTGAAACACGACCTGATC	489
	/db_xref="GI:145928"		125	IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr	144
	/translation="MLAFTRFLPAGLLVISHAFAMVODEHGTFTLEKTPQRIIVVLE		490	NTGCCGACGAGTGCCTCGCGGGTTTACATCGCTTGCAGCAAAATCGCCCGGTA	549
gene	1080..2078		145	IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle	164
	/gene="fecC"		550	CTGCTGCTTAAAGTCCCGCAACCAACCTAGCTGAAATTTGCAATTCGCGGTATCATC	609
CDS	/gene="fecC"		165	SerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAspLysLys	184
	/transl_table=11		610	GGCGAAATGGTGGGTAAAAGCGAGATGCGAGCGACTGGACCACTGGAACAACAATAAGAGAGG	669
	/protein_id="AAA23763.1"		185	IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal	204
	/db_xref="GI:145926"		670	ATGGCGAGTGGCGCCAGCCAG-----CTTCCCAAGGGACACGCGTGGCTTTGGCACA	723
	/translation="MTAIRKPVLLMGLPVAALIIIPWLSFCYSALIPVSGADATRAILL		205	AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer	224
	PGLTPALVQLNQLPRSLVAVLIGASLALAGTLTQTLTHNPMSPSLLGINSAA		724	TCAACGGAACAGCAATTCACCTGCATCTCAGGAGACCTGGACCGCGACGCTGGCC	783
gene	2075..3031		225	GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu	244
	/gene="fecD"		784	TCTCTGGGGCTG-----AACGTTCCCGTGGCGATG	813
CDS	/transl_table=11		245	LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg	263
	/protein_id="AAA23764.1"		814	CGGGTGGTCCATCGCGCTCCATCGGCTGAGCAACTGCTGGCGGTCAATCTCTGCTGG	873
	/db_xref="GI:145927"		264	MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGluLys	283
	/translation="MKTALVITFLALAGCALLSLHMGVIPVFWPWRALLTDWQAGREHY		874	CTGCTGGTGGCCACTATCGC-----GAAGAGAGCATTTGTTAAACGCTGGCAACA	924
	PGLTPALVQLNQLPRSLVAVLIGASLALAGTLTQTLTHNPMSPSLLGINSAA		284	AspProValTyrLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg	303
	PGLTPALVQLNQLPRSLVAVLIGASLALAGTLTQTLTHNPMSPSLLGINSAA		925	GATCCGCTCTGGCAGATGTTAAACCGCGCGAGCAAGCAGCAGGTTGCTTCGTCGACAGT	984
gene	3032..3799		304	AspLeuTyrAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuVal	323
	/gene="fecE"		985	AACACCTGGCGCGGATCGCGCGGTATTTTGTCTCAGAGCGGTATTGCGCTGACACGGTA	1044
CDS	/transl_table=11		324	GluLeu 325	
	/protein_id="AAA23765.1"		1045	AAATC 1050	
	/db_xref="GI:145928"		RESULT 34		
	/translation="MTURTENLTVSYGTDKVLNDVLSLPTGKITALIGPNCCKSTL		AF326777	66714 bp	DNA linear BCT 18-JUN-2002
	PGLTPALVQLNQLPRSLVAVLIGASLALAGTLTQTLTHNPMSPSLLGINSAA		AF326777	Shigella flexneri 2a	SRL pathogenicity island, complete sequence.
	PGLTPALVQLNQLPRSLVAVLIGASLALAGTLTQTLTHNPMSPSLLGINSAA		AF326777	Shigella flexneri 2a	
	PGLTPALVQLNQLPRSLVAVLIGASLALAGTLTQTLTHNPMSPSLLGINSAA		AF326777.2	GI:21450881	
	PGLTPALVQLNQLPRSLVAVLIGASLALAGTLTQTLTHNPMSPSLLGINSAA		Shigella flexneri 2a		
	PGLTPALVQLNQLPRSLVAVLIGASLALAGTLTQTLTHNPMSPSLLGINSAA		Shigella flexneri 2a		
	PGLTPALVQLNQLPRSLVAVLIGASLALAGTLTQTLTHNPMSPSLLGINSAA		Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;		
	PGLTPALVQLNQLPRSLVAVLIGASLALAGTLTQTLTHNPMSPSLLGINSAA		Enterobacteriaceae; Shigella.		
	PGLTPALVQLNQLPRSLVAVLIGASLALAGTLTQTLTHNPMSPSLLGINSAA		1 (bases 1 to 66714)		
	PGLTPALVQLNQLPRSLVAVLIGASLALAGTLTQTLTHNPMSPSLLGINSAA		Luck,S.N., Turner,S.A., Rajakumar,K., Sakellariis,H. and Adler,B.		
	PGLTPALVQLNQLPRSLVAVLIGASLALAGTLTQTLTHNPMSPSLLGINSAA		Ferric citrate transport system (fec) of Shigella flexneri 2a		
	PGLTPALVQLNQLPRSLVAVLIGASLALAGTLTQTLTHNPMSPSLLGINSAA		YSH6000 is encoded on a novel pathogenicity island carrying		
	PGLTPALVQLNQLPRSLVAVLIGASLALAGTLTQTLTHNPMSPSLLGINSAA		multiple antibiotic resistance genes		
	PGLTPALVQLNQLPRSLVAVLIGASLALAGTLTQTLTHNPMSPSLLGINSAA		Infect. Immun. 69 (10), 6012-6021 (2001)		
	PGLTPALVQLNQLPRSLVAVLIGASLALAGTLTQTLTHNPMSPSLLGINSAA		21437601		
	PGLTPALVQLNQLPRSLVAVLIGASLALAGTLTQTLTHNPMSPSLLGINSAA		11553538		
	PGLTPALVQLNQLPRSLVAVLIGASLALAGTLTQTLTHNPMSPSLLGINSAA		2 (bases 1 to 66714)		
US-10-724-972A-6352 (1-335) x ECOFECBCDE (1-4842)					
Qy	45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu		64		
	250 GTTCAGGACGACACGCGCGTTTACACTCGAAAAACCGCCACCGGATTTGGTGCTG		309		
Db					
Qy	65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp		84		
Db					
ORIGIN					
Alignment Scores:					
Pred. No.:	3,36e-25		Length:	4842	
	463.00		Matches:	100	
Score:	57.45%		Conservative:	62	
	35.4%		Mismatches:	104	
Best Local Similarity:	27.40%		Indels:	16	
			Gaps:	4	
Query Match:					


```

AUTHORS Luck,S.N., Turner,S.A. and Rajakumar,K.
TITLE Direct Submission
JOURNAL Submitted (06-DEC-2000) Microbiology, Monash University, VIC 3800,
Australia
REFERENCE 3 (bases 1 to 66714)
AUTHORS Luck,S.N., Turner,S.A. and Rajakumar,K.
TITLE Direct Submission
JOURNAL Submitted (17-MAY-2002) Microbiology, Monash University, VIC 3800,
Australia
REMARK Amino acid sequence updated by submitter
REFERENCE 4 (bases 1 to 66714)
AUTHORS Luck,S.N., Turner,S.A. and Rajakumar,K.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-2002) Microbiology, Monash University, VIC 3800,
Australia
REMARK Sequence update by submitter
COMMENT On Jun 18, 2002 this sequence version replaced gi:15808696.
FEATURES
     source
         1..66714
             /organism="Shigella flexneri 2a"
             /mol_type="genomic DNA"
             /db_xref="taxon:42897"
             /note="corresponds to Escherichia coli min 23"
         458..66714
             /note="SRL pathogenicity island"
         619..1836
             /gene="int"
         619..1836
             /gene="int"
         /note="member of the P4 bacteriophage integrase family"
         /codon_start=1
         /transl_table=11
         /product="integrase"
         /protein_id="AAL08426.1"
         /db_xref="GI:15808697"
         /translation="NAVLTDTKARHKDDPKLPKHGGITGLTLHPSSVKGKGVKRY
VSPVTKRRNAGLGTPEVSAEARTARIMREQLAAGDDPLEIKKAESEKVITPFA
DAARKVAELSGPENPKHVRKQSLTENIYAPQLGAKTLDLSITAADVAETLRPVWLT
LSEATSRVKQPIHVVMQWAGFCVANPVDHLLPQOTGRDEHOPAMPWOLPL
FVATSVYDEPNVTRALLMVLVTATPSGEARGWRWAEIDPFKRWITIPAEKARL
QHRVPLSQAIYILNIGLHDELVPFSPKQIILSDMWLTSFLRKKAASDIPGRVA
TAHGRSTFRWCSEQGSYRDLAERLAHTLUNKVEAAHYHRTDLLEQRVPMQAWADY
VMSLAKNTKGEFA"
         complement(2006..3535)
         /note="orf2; similar to YffJ on the cryptic P4 prophage
CP4-57 of Escherichia coli K-12"
         /codon_start=1
         /evidence="not experimental"
         /transl_table=11
         /product="unknown"
         /protein_id="AAL08427.1"
         /db_xref="GI:15808698"
         /translation="WMYTKLSLPDLTKLPDRQALLDEMHEETGIRREILLVMLT
VMAASVDTHEVLSGGORTSLQIFMCLSSASGSGKTSACAKLIAPVHETEELHQA
YIDDKNDYMMEMWTTDKILLERYKKEIERSPENAAAAAEKCIANKPEPPVQOV
LIVNDTPGEGIALKSLSPILLSDEGGTLDKRFKRSALYNTLWSGQPVVERAS
RPGRLANDVRLTILITOPVTFNKFFTLTGQIDRNGFLARLLCEGAKIMVTEPH
TATPVVTOCASCFKRSKSGSIOBDSLRSRERAKSKERICMTLSNAAGALERPHK
ENMDVTQNPWMTGFEDIIIVKREQVRIIAALLEKDPDSTVITLESNTSAIYLIEF
YFKHLIYKLSREISPAEKLDKWLQRIITTAGYIFQKSYIQAYAPYALKKCVLDE
ALDLAERQKIRIDNVLVYIGNTITPSELAKKLNIAPADAGVFCIDHQNKILYHNR
L"
         complement(3507..3812)
         /note="orf3; similar to Alpha on the cryptic P4 prophage
CP4-57 of Escherichia coli K-12"
         /codon_start=1
         /transl_table=11
         /product="putative transcriptional regulator"
         /protein_id="AAL08428.1"
         /db_xref="GI:15808699"
         /translation="MHMENNMSQTGLYTPSRILRMAELTRLGLISRSTIYVKINPA
SKYHDTFPPIRLGSGVGRASADEWMLMLTAPAWSEPNKLELNDDDDTITP"
         3868..5730
     CDS
         10422..10905
             /note="orf6; similar to YffJ on the cryptic P4 prophage
CP4-57 of Escherichia coli K-12"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08430.1"
             /db_xref="GI:15808701"
             /translation="MYNANPNYKMDFMILKDVNEHMDGMFORFSLKLPFRIDFAVRKD
TPSFQSGKSHGCMCEIYRLLSSETQTLAGYVWMEYTPDKGLHIHFVGLYDQGRHNS
YQISQLGDIWRRITEGGYFHLCKRAKQYIPVRIHVHYSKSAVDLRLYALSILAK
ODOKEGHIIIGRSRLPEKSNRGRPDITELWGLMSELS"
             8449..8745
             /note="orf6; similar to YffJ on the cryptic P4 prophage
CP4-57 of Escherichia coli K-12"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08431.1"
             /db_xref="GI:15808702"
             /translation="MSLSRLPFFFNKDAYYHLGNVRSYDRLGLITGAWKSALGLELED
HQGLVHPFANGRYILDINDPNFPQYQRLVRLDLYTLKLDTKASQRNFGVSGF"
             complement(9112..10421)
             /note="IS629-like element"
             10422..10905
             /note="remnant of IS200-like element"
             complement(10980..11918)
             /note="orf7; similar to a LysR-like transcriptional
regulator"
             /codon_start=1
             /transl_table=11
             /product="putative transcriptional regulator"
             /protein_id="AAL08432.2"
             /db_xref="GI:21450882"
             /translation="MKNIEWKLYDFLTLEACHFSQAQKERNLSQAFSRRIKALEA
AIGVLFDRITTLQTEBGLFSQTRSLQCLELNGSLGQSLGVPNRIKAAAH
SLSLVLPKLVHSLTAYGFEVHVEAIDVVQAVNTLRKESDFIISFRDMLQSPF
CCLKFSESELPVCAADAQGHVDFITQPVPLNLTATSYMGRLVNVRHLAEVGGITG
TRIFISMSSELKKNALNGYIAWLPWISVDELQTKRLICLDAKLTVPQIAYIRM
NTRLNRTAENLWRLQLQEHMPDDLIQIISMEEPARN"
             12299..12988
             /note="orf8"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08433.1"
             /db_xref="GI:15808704"
             /translation="MGPAATVDLNFKNVFTFAAQCOQEHIPLIISIPDIPORTDALM
HAGRSPLPVNMRDYMHLKEDAGAEICVPCNTHYHFKELKADCHIDLSIVETVINE
RACGKTRIGLLATNATLYMGLYQKIGIESITCTVSPDADQKQKMWESIYSLKAGNIAH
AOKLMNEQAEILFSRGAELVLGCTEIPVILAQVKEQSPRYIDSTASLVRAGIKWTE
NRVGKHLITQ"
             13054..14352
     CDS
         7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDAFI
EKPLIDPVYKEGEIOYDANNFTLMRVIKYCNMLEKISTKPIAPPYITGLDDVEK
VLKINILIDKEYVYDGTLAEVIMENKVLSSRERKETWIGLFTGSKKYTLQOCVEKL
GVLVHVKSPVDEIKNVMWLYKGAENRRNRMYDALTIICEDDIRNPPELS"
             7104..7697
             /note="orf5"
             /codon_start=1
             /evidence="not experimental"
             /transl_table=11
             /product="unknown"
             /protein_id="AAL08429.1"
             /db_xref="GI:15808700"
             /translation="MQTNRAQHMRTFTNMLYDICTVLGLFKEGENPAHKRKNSTNPFEM
HOKFMDQRIEISRIIDAGVFSQEQRRRIIYARYEHFYMMNSYPVHSTLKPFEFLRSY
CURTFGVIFVDMVNTYTPENDSAFYHIYNFLOKSYCPCLDHAOTESDEAAVKRYL
REYLAELFKENKDFHENGKLYALGKYTGITRKONGKSKLMQOYIKAIKNEYKDYRE
KKLDKDEKELVRNIDKFYNAPVSLVLDIQRKTILOSAYVLRVLVREGIWHGL
YGAAYLYDFTSFDTTPYAKLLEMFYKFSAGETLSRYSVSLDDKQSEYISRLKD
LVFNINDKNGCDDAVLKLIISYFGQLQNEAVHTSCYETLAVICILNKNKINDVLOH
YDMMERKGLFGELPSGVVRGALSULDALEVKVNRKNIKGLSYFLWLDHVKAYQDA
```


/note="orf9; similar to DcuA in Escherichia coli K-12"
/codon_start=1
/transl_table=1
/product="putative anaerobic decarboxylate transporter"
/protein_id="AAL08434.1"
/db_xref="GI:15808705"
/translation="MWELIIVLGAIFGIRGGIGIGLGGGLGAILTLTGFLPMG
SPVDVILLIMTVVAASALQAAGMDYLVVRVAGNPMRNPXYINIIPIITWSMTIM
AGTGFIVFTSLPIAETSAKESGIRPRTLAGSVASQVALSGSPISAANAAMLTINEV
NGVSFTQMSVCLPTSFSAAMVAAPFASQGGELQDDVYLERLQGLGVKVENNSSI
KPCATLSGLFLLATVAIVLAAPQPLRPGFDIGKPMETRDIIICMLSAACLMVVIC
KATADIIITSGFAGMSLAVILGIVTGTTFIDAHLEIKDIDAGDILQAVPMLIAL
VIFGTCALYSQAGTPLIPLAVALDVPTMAILASVAVTGTVFLPTPTSLAAMEF
DTGTRVGVKYLNVHHPMLPGUGGIIAGVALGVFIAPMII"
/note="remanant of 15600-like element; central region
deleted"
repeat_region
complement(15323..16090)
/note="ISI element"
/complement(16171..16962)
gene
/genes="aadA1"
CDS
/genes="aadA1"
/note="confers streptomycin and spectinomycin resistance;
similar to AadA1 from Tn21 on plasmid R100"
/codon_start=1
/transl_table=1
/product="streptomycin adenyltransferase AadA1"
/protein_id="AAL08435.1"
/db_xref="GI:15808706"
/translation="MREVVIAEYSTQSEVVGVIETRLHLEPTLLAVHLVSGAVDGLKP
HSDIDLVTVTRVEDTERRALLNDLSETASPGESILRAVEVTVVHDDILIPWRYP
AKRLOFGKWRNDILAGIFEPATIDILAILLKAREHSVALVGPAAELFDFVPBQ
DLPEKLETITLWNSPDWDXDNRNVDTLSXIWXSXTVXIAPDKVADWAMERLPA
QYQPVILEARQXYLQBEDRLASRADQLSEFVHVYVKGKITKVGK"
complement(17075..17905)
gene
Alignment Scores:
Pred. No.: 7,81e-24 Length: 66714
Score: 463.00 Matches: 100
Percent Similarity: 57.45% Conservative: 62
Best Local Similarity: 35.46% Mismatches: 104
Query Match: 27.40% Indels: 16
DB: 1 Gaps: 4
US-10-724-972A-6352 (1-335) x AF326777 (1-66714)
QY 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 39269 GTTCAGGACGACACGGCAGCTTTACCTCGAATAAACGCCACACGAGTTGTGTGTG 39210
QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyLeuAlaLeu 84
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 39209 GAACCTCTGTTCCGCGATGGCTGCGCCGCGGAGCTGAGCCGCGATTTGCGCGAC 39150
QY 85 AspAsnLysLysAsnArgIleLysProLysArgAspLysIleGlyLysTyrThrSer 104
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 39149 GATAACGATGCAAAACGATCTGCGCCGAGCGTGGCGACCTGAAACCGTGGCAGTCC 39090
QY 105 ValGlyThrArgLysGlnProAsnLeuGluGlyLeuSerLysLysLysProAspLeuLeu 124
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 39089 GTTCGAAACCGCGCGAGCGCTGGAAGCCATTTCCGCTCTGAAACACACCTGATC 39030
QY 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLysLysLysLeuAlaProThr 144
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 39029 ATTGCCGACGACGATGCCATCGCGGGTTTACATCGCTTGCAGCAATCGCGCGGTA 38970
QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 38969 CTGCTGTTAAGTCCCGCAACGAAACCTACGCTGAAATTTTGCATCTCGCGGTATCATC 38910
QY 165 SerLysAlaLeuGlyLysGluGluGlyLysLysLysLysLysLysLysLysLysLys 184
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

Db 38909 GCGGAATGTGGGTAAAGACGACAGATGCGAGGACAGCTCTGGAACAACATAAGAGAGG 38850
QY 185 IleGluGlyTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 38849 ATGGCGCAGTGGCCAGCCAG-----CTTCCAAAGGACACGCGTGGCTTTGGCACA 38796
QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 38795 TCAGCGAAACAGCAATTCACCTGCATACCTAGGAGACCTGGACCGCGAGCGTGTGGCC 38736
QY 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLysSerLysTyrLeu 244
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 38735 TCTCTGGGGCTG-----AACGTTCCCGCTGCGATG 38706
QY 245 LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 38705 CGGGTTCGTCATGCGGTCCTCGGCAACTGCTGGCGGTCAATCTCTGCTGG 38646
QY 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGluLys 283
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 38645 CTGCTGTTGCCCACTATCGC-----GAAGAGAGCATTTGTTAAACCGCTGCAACAA 38595
QY 284 AspProValTrrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg 303
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 38594 GATCCGCTCTGGCAGATGTTAACCGCGCCGACGAGCAGGTTGCTTCGTCGACAGT 38535
QY 304 AspleuTrrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuVal 323
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 38534 AACACCTGGCGCGGATGCGCGGTATTTTGTGCAGAGGTATTCGCGTACACGGTA 38475
QY 324 GluLeu 325
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 38474 AAAATC 38469
RESULT 35
AX370197/c 10244 bp DNA linear PAT 16-FEB-2002
LOCUS
DEFINITION Sequence 14 from Patent WO0170776.
ACCESSION AX370197
VERSION AX370197.1 GI:18857380
KEYWORDS Escherichia coli
SOURCE Escherichia coli
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE
AUTHORS Levy, S.B., Barbosa, T.M. and Alekshun, M.N.
TITLE Nmr compositions and their methods of use
JOURNAL Patent: WO 0170776-A 14 27-SEP-2001;
TRUSTEES OF TUFTS COLLEGE (US)
FEATURES
Location/Qualifiers
source
1. 10244
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"
ORIGIN
Alignment Scores:
Pred. No.: 1.39e-24 Length: 10244
Score: 460.00 Matches: 100
Percent Similarity: 57.45% Conservative: 62
Best Local Similarity: 35.46% Mismatches: 104
Query Match: 27.22% Indels: 16
DB: 6 Gaps: 4
US-10-724-972A-6352 (1-335) x AX370197 (1-10244)
QY 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 7735 GTTCAGGACGACACGCGCAGTTTACCTCGAATAAACGCCACACGAGTTGTGTGTG 7676
QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyLeuAlaLeu 84
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

Db 7675 GAACCTCGTTGCGCGATCGCTGGCCGCGTGGACGTCATCCCGATCGGTATTGCGGAC 7616
Qy 85 AspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSer 104
Db 7615 GATAACAGTGAACAAACGATCTGCGGAGTGGCGACCTGAAACCGTGGCAGTCC 7556
Qy 105 ValGlyThrArgLysGlnProAsnLeuGluLysSerLysLysLeuLysProAspLeuIle 124
Db 7555 GTCCGACGCGCGCGACGAGCTGGAAGCCATTCGCGTCTGGAACACAGACCTGATC 7496
Qy 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144
Db 7495 ATTGCCGACAGCAGTCCGCGGTTTACATCCCTTGCAGCAATTCGCGCCGGTA 7436
Qy 145 IleGluLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164
Db 7435 CTGCTGTTAAGTCCCGCAACGAACTACGCTGAAATTTGCAATTCGCGCTATCATC 7376
Qy 165 SerLysAlaLeuGlyLysGluGluGlyLysArgLeuGluGluHisAspLysLys 184
Db 7375 GGCGAATGTTGGTAAAGACGAGATCGCGCAGTCTGGAACACATTAAGAGAGG 7316
Qy 185 IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204
Db 7315 ATGGCGAGTGGCCAGCCAG-----CTTCCCAAGGACACGCGTGGCTTTGGCACA 7262
Qy 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224
Db 7261 TCACGCGACAGCAATTCAACCTGCATCTCAGGACAGCTGGACCGCGCTGGCC 7202
Qy 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu 244
Db 7201 TCTCTGGGCTG-----AACCTTCCCGCTGGATG 7172
Qy 245 LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263
Db 7171 GCGGGTGGCTCCATCGCGTCCATCGCGCTGGAGCAACTCTGCGGTCAATCTGCGTG 7112
Qy 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLysLysGluLeuGluLys 283
Db 7111 CTGCTGTTGCCACTATCGC-----GAAGAGAGCATTTGTAACGCTGGCAACA 7061
Qy 284 AspProValTyrLysLysAsnAlaValLysAsnGlnArgValAspIleLeuAspArg 303
Db 7060 GATCCGCTCTGCGCATGTATACCGCGCGAGAGCAGCAGGTGCTTCGGTGCACAGT 7001
Qy 304 AspLeuTyrAlaArgSerArgGlyLeuIleSerSerGluLysMetAlaLysGluLeuVal 323
Db 7000 AACACCTGGCGCGATGCGCGTATTTTGTGCAGAGCGTATTGCGCGTGCACACGTA 6941
Qy 324 GluLeu 325
Db 6940 AAAATC 6935

RESULT 36
U00096_45/c
WPCOMMENT

Sequence split into 47 fragments LOCUS U00096 Accession U00096

Fragment Name	Begin	End
U00096_00	1	110000
U00096_01	100001	210000
U00096_02	200001	310000
U00096_03	300001	410000
U00096_04	400001	510000
U00096_05	500001	610000
U00096_06	600001	710000
U00096_07	700001	810000
U00096_08	800001	910000
U00096_09	900001	1010000
U00096_10	1000001	1110000
U00096_11	1100001	1210000
U00096_12	1200001	1310000
U00096_13	1300001	1410000

U00096_14 1400001 1510000
U00096_15 1500001 1610000
U00096_16 1600001 1710000
U00096_17 1700001 1810000
U00096_18 1800001 1910000
U00096_19 1900001 2010000
U00096_20 2000001 2110000
U00096_21 2100001 2210000
U00096_22 2200001 2310000
U00096_23 2300001 2410000
U00096_24 2400001 2510000
U00096_25 2500001 2610000
U00096_26 2600001 2710000
U00096_27 2700001 2810000
U00096_28 2800001 2910000
U00096_29 2900001 3010000
U00096_30 3000001 3110000
U00096_31 3100001 3210000
U00096_32 3200001 3310000
U00096_33 3300001 3410000
U00096_34 3400001 3510000
U00096_35 3500001 3610000
U00096_36 3600001 3710000
U00096_37 3700001 3810000
U00096_38 3800001 3910000
U00096_39 3900001 4010000
U00096_40 4000001 4110000
U00096_41 4100001 4210000
U00096_42 4200001 4310000
U00096_43 4300001 4410000
U00096_44 4400001 4510000
U00096_45 4500001 4610000
U00096_46 4600001 4639675
Continuation (46 of 47) of U00096 from base 4500001 (U00096 Escherichia coli K-12 MG165)
Alignment Scores:
Pred. No.: 2,39e-23 Length: 110000
Score: 460.00 Matches: 100
Percent Similarity: 57.45% Conservative: 62
Best Local Similarity: 35.46% Mismatches: 104
Query Match: 27.22% Indels: 16
DB: 1 Gaps: 4
US-10-724-972A-6352 (1-335) x U00096_45 (1-110000)
Qy 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64
Db 12262 GTTCAGGACGAACACGCGACGCTTTACACTCGAAAAACGCCACACGATTTGTGTCTG 12203
Qy 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84
Db 12202 GAATCTCGTTGCGCGATCGCTGCGCCCGCTGGACGTCATCCGATCGTATTGCGGAC 12143
Qy 85 AspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSer 104
Db 12142 GATAACGATGCACAAACGATCTCCCGAAGTGGTGGCGACCTGAAACCGTGGCAGTCC 12083
Qy 105 ValGlyThrArgLysGlnProAsnLeuGluLysSerLysLysLeuLysProAspLeuIle 124
Db 12082 GTCCGAACGCGCGCGACGCGGCTGGAGCCATTGCGCGCTCTGAAACACGACCTGATC 12023
Qy 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144
Db 12022 ATTGCCGACAGCAGTCGCCATTCGCGGGGTTTACATCGCTTGCAGCAATTCGCCCGTA 11963
Qy 145 IleGluLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164
Db 11962 CTGCTGCTTAAGTCCCGCAACGAACTACGCTGAAATTTGCAATTCGCGCTATCATC 11903
Qy 165 SerLysAlaLeuGlyLysGluGluGlyLysArgLeuGluGluHisAspLysLys 184
Db 11902 GGCGAATGTTGGTAAAGCGAGATGTCAGGACGCTCTGGAAACCAATAAGAGAGG 11843

Qy	185	IleGluGluTyrIleGluLeuThrMetAspLysGlnValLeuProAlaVal	204
Db	11842	ATGGCGCAGTGGCCAGCCAG-----CTTCCAAAGGACACGCGTGGCTTGGCACA	11789
Qy	205	AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer	224
Db	11788	TACGGGAACAGCAATTCAACCTGATCTACGAGACCTGGACCGCAGCGTGGCC	11729
Qy	225	GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu	244
Db	11728	TCTCTGGGGCTG-----AACGTTCCCGTGGCGATG	11699
Qy	245	LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg	263
Db	11698	CGCGTGGCGTTCATCGCGCTGACAACTGCGCGGTCAATCTCGCTGG	11639
Qy	264	MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLys	283
Db	11638	CTGCTGTTGCCCACTATCC-----GAAGAGAGCATTTGTTAAACGCTGCAACAA	11588
Qy	284	AspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg	303
Db	11587	GATCCGCTCTGGCAGATGTTAACCGCGCGCAGAGCAGGTTGCTCGGTGACAGT	11528
Qy	304	AspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuVal	323
Db	11527	AACACCTGGCGGGATGCGCGGTATTTTCTGCAGAGCGTATTGCGCGTGACACGTA	11468
Qy	324	GluLeu	325
Db	11467	AAATC	11462
LOCUS	ECOW93	338534 bp	DNA linear BCT 30-JAN-2001
DEFINITION	Escherichia coli K-12 chromosomal region from 92.8 to 00.1 minutes.		
ACCESSION	U14003		
VERSION	U14003.1	GI:1263172	
KEYWORDS	Escherichia coli		
SOURCE	Escherichia coli		
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.		
REFERENCE	1 (bases 1 to 338534)		
AUTHORS	Burland, V., Plunkett, G. III, Sofia, H. J., Daniels, D. L. and Blattner, F. R.		
TITLE	Analysis of the Escherichia coli genome VI: DNA sequence of the region from 92.8 through 100 minutes		
JOURNAL	Nucleic Acids Res. 23 (12), 2105-2119 (1995)		
MEDLINE	95334362		
PUBMED	7610040		
REFERENCE	2 (bases 1 to 338534)		
AUTHORS	Plunkett, G. I. I.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-AUG-1994) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA. Email: ecoliegenetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459		
COMMENT	On Apr 17, 1996 this sequence version replaced gi:536929. This sequence was determined as part of the E. coli Genome Project (Frederick R. Blattner, director) at the University of Wisconsin-Madison. Supported by award HG00301 from the NIH Human Genome Project. The entire sequence was independently determined from E. coli MG1655; overlaps with other sequence determinations are annotated. The start of this entry overlaps the end of the entry ECOW89 (U00006) by 1885 bp.		
FEATURES	Location/Qualifiers		
source	1. 338534		
	/organism="Escherichia coli"		
	/mol_type="genomic DNA"		
	/strain="K-12"		
	/sub_strain="MG1655"		
	/db_xref="taxon:562"		
	/note="This sequence comprises the following lambda clones: DD928(EC27-278), DD930(EC30-188), DD933(EC27-1320), DD935(EC30MM32), DD937(EC27-1070), DD941(EC17-142), DD945(EC17-8), DD947(EC24A-34), DD949(EC22-169), DD952(EC27-297), DD953(EC27-409), DD956(EC17-103), DD958(EC30MM1), DD960(EC30MM2), DD962(EC21-104), DD965(EC23A-40), DD968(EC30K660A-4pp), DD970(EC19-202), DD974(EC19-61), DD975(EC18-233), DD977(EC30MM7), DD980(EC27-1151), DD981(EC18-3), DD984(EC17-101), DD987(EC18-115), DD990(EC17-136), DD992(EC18-282), DD995(EC17-115), DD997(EC18-113); M13mp19 or Janus vectors were used for subcloning"		
	1. 1885		
	/notes="1885 bp overlap with end of GenBank Accession Number U00006 (ECOW89)"		
	complement(41..39)		
	/notes="ORF_F326 of GenBank Accession Number U00006 (ECOW89)"		
	/codon_start=1		
	/transl_table=11		
	/label=ORF_f132p		
	/protein_id="AAA96985.1"		
	/db_xref="GI:1263173"		
	/translation="MGFTTRVKSSEKPFNFALFWDKYGTFILAIIVAFGSLSP EYFLTNNTIQTIFVQSVTVLIGWGEFFAILVAGIDLSVGAIALSGMVTAKMLAGV DFLAAMIGVLVGGALGALNGCLVNWTLG"		
	277..397		
	/note="predicted bend of 75.10 degrees"		
	complement(377..1909)		
	/notes="ORF_f510"		
	/codon_start=1		
	/transl_table=11		
	/label=ORF_f510		
	/protein_id="AAA96986.1"		
	/db_xref="GI:536931"		
	/translation="MATPYISMAGIGKSGFVHALKSVNLTVTPGETHALLGENGAGK STLKVLGSHHEPTKGTITINYNKDLHKLAAQLGIYQSLVIDEVLVLENYL IGRHLTKKICGVNIIDWREVRVAMMLRVLGVDLDEKVANLSHKKQLEIAKTL MLDKAVIMDEPTSSLTKEVDYLFILMNLQRKEGTAVTISHKLAIRICDRTVM KDSSVSGIVSDVNDIVRLMVRQLONFAMKVNLSLAHETVFEVNVTSRDR KKVNDISFVCRGEIIFAGLVSGRTELNLCLGVDRKAGGEIRLAKGDI:SPRSPLD AVKGMAYITTESRRDNGFFNFISIAQNAISRLSKWCCCEVPIIPEIPRIGIDVGAKEIYK RELIALKCHSVNQITELSGNQOQVLISKWLCCCEVPIIPEIPRIGIDVGAKEIYK VMRLQADDGKVLMSSELPEITVCDRIAIVFCGRLTQILTNRDDMSSEIEIMAWALP QE"		
	809..931		
	/note="predicted bend of 74.75 degrees"		
	1768..1889		
	/note="predicted bend of 75.87 degrees"		
	complement(1983..2010)		
	/note="promoter matrix score of 55; putative; resembles flh-regulated promoter"		
	complement(1985..2021)		
	/standard_name="REP; repetitive extragenic palindromic element"		
	/note="contains 1 REP sequence"		
	complement(2036..2971)		
	/note="ORF_f311"		
	/codon_start=1		
	/transl_table=11		
	/label=ORF_f311		
	/protein_id="AAA96987.1"		
	/db_xref="GI:536932"		
	/translation="MNKYLKYPFSGTLVGLMLSTSAFAAEYAVVLKTLSPNPFVDMKK GIEDEATLGVSDVIFASPSGDFQSLQFEDLSNKYKGIAPLPSVNLVMPVAR AWKGIYLVNLDKIDMDNLKKAGNVFAFTDNNVAGKAGASFIIDKLAGEGEVA IIEGKAGNSGEARRNGATEAFKASQIKLASQPADWDRIKALDVATNVLRQNPFIK IIEYCANDTMAMGVAQAVANAGTKGLVVGVTGDIPEARQWVEAGQMTATVQNPADIG ATGLKLVDAEKSQKVPILDKAPEFKLVDSILVTQ"		
	3010..3121		
	/note="predicted bend of 73.22 degrees"		
	complement(3030..3953)		
	/misc_feature		
	misc_structure		
	CDS		
	misc_structure		
	CDS		
	promoter		
	repeat_region		
	CDS		
	misc_structure		
	CDS		

```

/note="ORF_f307"
/codon_start=1
/transl_table=1
/label=ORF_f307
/protein_id="AAA96988.1"
/db_xref="GI:536933"
/translation="MASSIQQENVMQSBFDSALPNIGLAPYLVRMKQSGWTENESR
IVEMLLKPGNLSKAPAIKDVAAELAVSEAMIKVSKLIGSGERNLBSALEDFPSQE
QVLPSLAFDEAPQDVNVKFNITLRTIMEGQSIYVNDIHRARFYQARQDLIGA
GGSNAICADVQHEFLRGVRCQAYPDHIMMGSALIQEGDVLVWTHSGRTSDVKAA
VELAKNGAKILCITHSYHPIAKLADYIICSPAPETPLGRNASARILQULIDAFF
VSAQLNIEQANIMQKTGAIVDFPFSFGALK"
/promoter
misc_structure
4151..4251
/note="predicted bend of 61; putative"
4279..4728
/gene="y7cA"
4279..4728
/gene="y7cA"
/codon_start=1
/transl_table=1
/label=ORF_0149
/protein_id="AAA96989.1"
/db_xref="GI:536934"
/translation="MKKIAFGCDHVGFIKHEIVAHLVGVGVIEIDKGTWSEBTDYP
HYASOVAVAGCEVGGILICGTGVSIAANKFAGIRAVVCSSEPSYSLSRQHNDF
NVLAGRSRVGELAKMIVDAMUGAQYEGGRQQRVEAITAEQRN"
complement(4520..16191)
/note="corresponds to GenBank Accession Number D90227
Number J05260 (SCOPHNAQ) is from E. coli B, not K-12"
complement(4739..5146)
/gene="phnQ"
complement(4739..5146)
/gene="phnQ"
/note="annotated in GenBank Accession Number J05260
(SCOPHNAQ), but probably not a gene"
/codon_start=1
/transl_table=1
/protein_id="AAA96990.1"
/db_xref="GI:536935"
/translation="MPAPVPTSSPILIFHSPILFLOLTLIKLALAIQVLPAAIFKL
AAAVILSAPTVAHFTLPAVVARIPIVFTVRVSVINRLNIANGVVDNIGLGRNVD
SVSRVKNRSGKRFISQPKETLSLIITHPCK"
complement(5272..6030)
/gene="phnP"
complement(5272..6030)
/gene="phnP"
/codon_start=1
/transl_table=1
/protein_id="AAA96991.1"
/db_xref="GI:536936"
/translation="WSLTLTLTGTGGAGVPWGCCECAACARARRSPQYRRQPCGVV
KFNDAILIDAGLHDLADRWSPGSFQQLLTHYMDHVQGLFPLRWGDDPIYVGP
DEQCDLDFKHPLDLSHTVEPFFVFDLQGLVTPILNHSKLTFFGLLETAHSRA
WLSATGPEKTLKFLRNQPVVMDCSHPPRADAPRNHCDLNTVLALANQVIRSPV
ILTHISHQFDALMENALPSGFVGFDFGMEIGVA"
5452..5586
/note="predicted bend of 75.60 degrees"
complement(6032..6466)
/gene="phnO"
complement(6032..6466)
/gene="phnO"
/codon_start=1
/transl_table=1
/protein_id="AAA96992.1"
/db_xref="GI:536937"
/translation="MPACELRPATQYTDVAYILCELKQAEFDHAFRVGFNANLRD
PNMRYHLADJGVGVGIMGLHQLPHLHVNNWIGIEIOELVMPQARGNLVSGSKLLAW
EERQAQAEINTELSNVKRDHAFVYLRGEYQSHFRFTAL"
complement(6453..7010)
/gene="phnN"

```

```

CDS
complement(6453..7010)
/gene="phnN"
/note="HisP-like nucleotide binding protein (ATP-binding
protein)"
/codon_start=1
/transl_table=1
/protein_id="AAA96993.1"
/db_xref="GI:536938"
/translation="MMGKLIWMPSSGSKDSLLAEILREOTOLLVAHYRITRDASA
GSNNHIALSQEFPTAGQNLALLSHANGLIYGVGVIEDLWHLHAGFDVLVNGRAHL
PQARARYQSALLPVLQVSPFELIRQRLNRENGRENASEINARLARAARYTPQDCHTLAN
DGSLSQSVDTLLTIHQEKHACL"
complement(7010..8146)
/gene="phnM"
complement(7010..8146)

Alignment Scores:
Pred. No.: 9,21e-23 Length: 338534
Score: 460.00 Matches: 100
Percent Similarity: 57.45% Conservative: 62
Best Local Similarity: 35.46% Mismatches: 104
Query Match: 27.22% Indels: 16
DB: 1 Gaps: 4

US-10-724-972A-6352 (1-335) x ECUW93 (1-338534)
QY 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64
Db 205073 GTTCAGGACGACACGCGACGTTTACATCGAAAAACCCACACGCGATTGGTGTG 205014
QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84
Db 205013 GAATCTCTGTCGCGATGCGTGGCCGCGGTGGAGCTCATCCCGATCGGTATTGCG 204954
QY 85 AspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysThrSer 104
Db 204953 GATAACGATCGAAAAACGATCTCTCCGGAAGTGGTGGCCACCTGAAACGCGGAGTCC 204894
QY 105 ValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIle 124
Db 204893 GTCCGACGCGCGCGCGAGCGCTGGAAGCATTGCGCTCTGAAACACAGACCTGATC 204834
QY 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144
Db 204833 ATTGCGACAGCAGTGGCCATCGCGGGGTTTACATCGCTTGCAGCAATCGCGCGGTA 204774
QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164
Db 204773 CTGCTGCTTAAGTCCCGCAACGAAACCTACGCTGAAATTTGCAATCTCGCGCTATCATC 204714
QY 165 SerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAspLysLys 184
Db 204713 GGCGAAATGGTGGGTAAAGCAGAGATGAGCGACGCTCTGGAACAACATAAGAGAGG 204654
QY 185 IleGluGluTyrLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204
Db 204653 ATGCGCGAGTGGCGCCAGCCAG-----CTTCCCAAGAGGACACGCGTGGCTTTGGCACA 204600
QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224
Db 204599 TCACGCAACAGCAATTCACCTGCTACTACAGAGACCTGACCGGCGAGCGTGTGCGCC 204540
QY 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu 244
Db 204539 TCTCTGGGGCTG-----AACGTTCCCGCTGCGATG 204510
QY 245 LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263
Db 204509 CGCGGTGCGTCCATGCGTCCATCGCGCTGAGCACTGTGCGCGGTCAATCTCTCGCTGG 204450
QY 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLys 283
Db 204449 CTGCTGTTGCCCACTATCGC-----GAAGAGAGCATTTGTTAAACGCTGGCAACA 204399

```

QY	284	AppProValTrrpLysLeuAenAlaValLysAenGlnArgValAspIleLeuAapArg 303	CDS	/locus_tag="PSPT00589" complement (1390..2199) /gene="hopPtoC"
Db	204398	GATCCGCTGGCAGATGTTAAACCGCCGACAGACAGAGTTCCTTCGGCTGCACAGT 204339		/locus_tag="PSPT00589" /note="identified by match to TIGR protein family HMM TIGR01586"
QY	304	AppLeuTrrpAlaArgSerGlyLeuIleSerSerGluGluMetAlaLysGluLeuVal 323		/codon_start=1 /transl_table=11 /product="type III effector HopPtoC" /protein_id="AAO54131.1" /db_xref="GI:28851053"
Db	204338	AACACCTGGCGCGGATGTCGGGTATTTTTCGTGCAGACGATATTCCTGCACACGGTA 204279		/translation="WTIVSGHIGKHPSLTTVOAGSSASVENOMPDPAQFSQGRWKKLP TQLSITLARFDODICTNNHGISQAMCFGLSLSWINMIHAGKDHWTPVASAERMFEL GSFEGVHARTVNFYRTEHKFLMEQASANPGVSGMAGTSELQAAELKGLKLPV LEDKNSGLPFLACKSGRQVSTDAALSSICDAIVENKRGVMVIYSIEIHAULGS VSDGKRATLPDENLGEFHTSKALADTIENISSADGLPLIGVQVFASKIH"
QY	324	GluLeu 325		2167..2310 /locus_tag="PSPT00590"
Db	204278	AAATC 204273		2167..2310 /locus_tag="PSPT00590" /note="identified by Glimmer2; putative"
RESULT 38	AE016858/c	Pseudomonas syringae pv. tomato str. DC3000 section 3 of 21 of the complete genome	gene	/codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAO54132.1" /db_xref="GI:28851054"
LOCUS	DEFINITION	Pseudomonas syringae pv. tomato str. DC3000	CDS	/translation="MFSDVSRHDCHTIFSALFKMRNWFHFRFSSLTFLKMGQPVFLG FGT"
ACCESSION	VERSION	AE016858.1 GI:28851051		2374..3474 /locus_tag="PSPT00591"
KEYWORDS	SOURCE	Pseudomonas syringae pv. tomato str. DC3000		/locus_tag="PSPT00591" /note="similar to GB:U10552, and PID:504482; identified by sequence similarity; putative"
ORGANISM	REFERENCE	1 (bases 1 to 312839) Buell, C.R., Joardar, V., Lindeberg, M., Selengut, J., Paulsen, I.T., Gwinn, M.L., Dodson, R.J., Deboy, R.T., Durkin, A.S., Kolonay, J.F., Madupu, R., Daugherty, S., Brinkac, L., Beanan, M.J., Haft, D.H., Nelson, W.C., Davidson, T., Zafar, N., Zhou, L., Liu, J., Yuan, Q., Khoury, H., Fedorova, N., Tran, B., Russell, D., Berry, K., Uttterback, T., Van Aken, S.E., Feldblyum, T.V., D'Acenozo, M., Deng, W.L., Ramos, A.R., Alfano, J.R., Cartinhour, S., Chatterjee, A.K., Delaney, T.P., Lazarowitz, S.G., Martin, G.B., Schneider, D.J., Tang, X., Bender, C.L., White, O., Fraser, C.M. and Collmer, A.	gene	/transl_table=11 /product="impB/mucB/samB family protein" /protein_id="AAO54133.1" /db_xref="GI:28851055"
TITTLE	JOURNAL	Pseudomonas syringae pv. tomato DC3000		/translation="MILLTKLHCLSGPLVFPFPHINGDMSERIMTIIESMVPASEVYS IDEPADLTGMGNLTFEGRFIRSTILKCTGIPVGIGIAPTKTAKLADHTARLLAQ TGGVVDICDHKNWLVNRNVCSEVWGKKAHLEAMNIRSLAKADARTLRT FSVVIETARELAGTSCLESEADPPKQECSSRFQRLTAIPIKEAVATYQRAA EKLRAQNSCLKKRVSI RTQMFNPDEPKYANGAMIELPYPTNDVRLMTKGTAEAVNRL FRPGYKSKAEVLILLDLRQGEFTDDILFAASQPAAEKVMGVLDLEINARMGRTLRIG SVPTNPEWAMRRDMSQSYTTRLDQLWTVRSE"
PUBMED	REFERENCE	Proc. Natl. Acad. Sci. U.S.A. 100 (18), 10181-10186 (2003)		3877..4476 /gene="trpG" /locus_tag="PSPT00592"
AUTHORS	FEATURES	1. 312839 Location/Qualifiers /organism="Pseudomonas syringae pv. tomato str. DC3000" /mol_type="genomic DNA" /strain="DC3000" /db_xref="taxon:223283" /note="pathovar: tomato" complement (82..738) /gene="hopPtoH" /locus_tag="PSPT00588" complement (82..738) /gene="hopPtoH" /locus_tag="PSPT00588" /codon_start=1 /transl_table=11 /product="type III effector HopPtoH" /protein_id="AAO54130.1" /db_xref="GI:28851052"	CDS	/protein_id="AAO54134.1" /db_xref="GI:28851056" /translation="MLLMDIYDSDFTYNNVOYLGBELGADVIRNDELITIAIEALNP ERIVVSPGCTPNEAGVSLVIEKHFGAKLPILGCLGHQSIGQAFGVDVVRQVMHG KTSVPHVEDGGVFPAGLNHPVAVTVHSLVAVKSDTLPECLEVTAWTALEDGSDVEIMGL RHKTUNVEGVQFHPESILTEQQHELPAFLKQSGGHRQG" 4498..5547 /gene="trpD" /locus_tag="PSPT00593"
gene	CDS	2 (bases 1 to 312839) Buell, R., Joardar, V., Khouri, H., Fedorova, N., Tran, B., Russell, D., Berry, K., Uttterback, T., Van Aken, S., Feldblyum, T., Gwinn, M., Dodson, R., Deboy, R., Durkin, A., Kolonay, J., Madupu, R., Daugherty, S., Brinkac, L., Beanan, M., Haft, D., Selengut, J., Nelson, W., Davidson, T., White, O., Fraser, C. and Collmer, A.		/codon_start=1 /transl_table=11 /product="anthranilate synthase, component II" /protein_id="AAO54135.1" /db_xref="GI:28851057"
TITTLE	JOURNAL	Direct Submission Submitted (03-MAR-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA	gene	/translation="WNKIALNRVNVQLDLSTDEMVDVMEITWGCTEAQICAFELMG
FEATURES	source	Location/Qualifiers 1. 312839 /organism="Pseudomonas syringae pv. tomato str. DC3000" /mol_type="genomic DNA" /strain="DC3000" /db_xref="taxon:223283" /note="pathovar: tomato" complement (82..738) /gene="hopPtoH" /locus_tag="PSPT00588" complement (82..738) /gene="hopPtoH" /locus_tag="PSPT00588" /codon_start=1 /transl_table=11 /product="type III effector HopPtoH" /protein_id="AAO54130.1" /db_xref="GI:28851052"	CDS	/product="anthranilate phosphoribosyltransferase" /protein_id="AAO54135.1" /db_xref="GI:28851057"
gene	CDS	SSLABGGORYNPRTGSGEELRAVGLDKYRYSLTTKKPSSENSIRASHGLPWRMKYRAHQ "		/translation="WNKIALNRVNVQLDLSTDEMVDVMEITWGCTEAQICAFELMG
gene	CDS	complement (1390..2199) /gene="hopPtoC"		

QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAsnAlaPheIleValThrIle 164

Search completed: November 9, 2005, 23:37:16
Job time : 7014 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 9, 2005, 10:23:21 ; Search time 641 Seconds

(without alignments)

3093.781 Million cell updates/sec

Title: US-10-724-972A-6352

Perfect score: 1690

Sequence: 1 GVSEVRLKLTSLVIGLLFVL.....EEMAKELVELSKDKSKDKNK 335

Scoring table:

BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delxext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-Q=/cgn2.1/USPTO spool_h/US10724972/runat_03112005_141859_29140/app_query.fasta_1.519

-DB=N Geneseq -QMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0

-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45

-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000

-USER=US10724972 @CNC 1 1 470 @runat_03112005_141859_29140 -NCPUP=6 -ICPU=3

-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_16Dec04:*

1: geneseqn1980s:*

2: geneseqn1990s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002as:*

7: geneseqn2002bs:*

8: geneseqn2003as:*

9: geneseqn2003bs:*

10: geneseqn2003cs:*

11: geneseqn2003ds:*

12: geneseqn2004as:*

13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1690	100.0	1008	13 ADS03285	Ads03285 Staphyloc
2	1671	98.9	993	8 ACA47084	Aca47084 Prokaryot
3	1242	73.5	1014	4 AAS54519	Aas54519 Staphyloc
4	1233	73.0	2981	8 ACC48532	Acc48532 Staphyloc
5	1232	72.9	984	8 ACA19774	Aca19774 Prokaryot

ALIGNMENTS

RESULT 1

ADS03285

ID ADS03285 standard; DNA; 1008 BP.

XX

AC ADS03285;

XX

DT 04-NOV-2004 (first entry)

XX

DE Staphylococcus epidermis polynucleotide seqid 2580.

XX

DE antibacterial; vaccine; antisense therapy; Staphylococcus epidermidis;

XX

DE recombinant expression vector; infection; computer readable medium;

XX

DE computer based system; gene; ds.

XX

OS Staphylococcus epidermidis.

XX

PN US2004147734-A1.

XX

FD 29-JUL-2004.

XX

PF 01-DEC-2003; 2003US-00724972.

XX

PR 08-NOV-1997; 97US-0064964P.

XX

PR 13-AUG-1998; 98US-00134001.

XX

PR 29-NOV-1999; 99US-00450969.

XX

PA (DOUC/) DOUCETTE-STAMM L.

XX

PA (BUSH/) BUSH D.

6 1214 71.8 1044 8 ACF74986
7 1181 69.9 927 4 AAS51822
8 726 43.0 957 4 AAS51612
9 726 43.0 960 4 AAS54387
10 724.5 42.9 960 8 ABT15015
11 724.5 42.9 960 8 ACA20079
12 724.5 42.9 975 8 ACF74415
13 724.5 42.9 2957 8 ACC48534
14 721 42.7 2247 2 AAT80398
15 721 42.7 2247 2 AAT83786
16 679.5 40.2 5718 12 ADF30765
17 644 38.1 525 8 ACA47237
18 516.5 30.6 897 8 ACA42836
19 510.5 30.2 801 6 ABK74775
20 476 28.2 909 10 ACF70890
21 476 28.2 25860 11 ADR20887
22 476 28.2 110000 10 ACF67367_39
23 476 28.2 110000 10 ACF65388_08
24 460 27.2 900 8 ACA32000_08
25 460 27.2 909 4 AAS52735
26 460 27.2 909 8 ACA32751
27 460 27.2 10244 4 AAS46238
28 454 26.9 891 8 ACA45366
29 405 24.0 990 8 ABT15032
30 405 24.0 999 2 AAZ22850
31 405 24.0 999 2 AAZ19889
32 405 24.0 999 10 ADF43555
33 405 24.0 2940 8 ACC48531
34 405 24.0 3775 2 AAV74549
35 391 23.1 972 8 ACA21545
36 380.5 22.5 944 6 ABK74770
37 375 22.2 2209 5 AAS71379
38 375 22.2 4392 5 AAS73038
39 355 21.0 668 2 AAV74666
40 350 20.7 2799 5 AAS77335
41 343 20.3 861 8 ACF74610
42 334 19.8 2115 8 ACF74915
43 333 19.7 242 2 AAV76867
44 331 19.6 945 5 AAH65340
45 331 19.6 349980 5 AAH68525

121 ProAspLeuIleIleAlaAspAsnArgHisLysGlyIleTyrLysAspLeuAsnLys 140
141 IleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAla 160
161 PheLysThrIleSerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGlu 180
181 HisAspLysLysIleGluLysLysGluLysGluLysGluLysMetAspLysAsnGlnLysVal 200
201 LeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGly 220
221 GlnPheLeuSerGlnLeuGlyPheLysGluLysGluLysSerAspValThrLysGlyLeu 240
241 SerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsn 260
261 ProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGlu 280
281 LeuGluLysAspProValTyrLysLysLeuAsnAlaValLysAsnGlnArgValAspIle 300
301 LeuAspArgLeuThrAlaArgSerArgGlyLeuLysSerSerGluGluMetAlaLys 320
321 GluLeuValGluLeuSerLysLysAspSerLysLysAspAsnLys 335
335 GAACTTGTGTAATTTATCTAAGAAAGATAGTAAAAAGAGTAATATAG 1005

RESULT 2
ACA47084
ID ACA47084 standard; DNA; 993 BP.
AC ACA47084;
AC ACA47084;
DT 19-JUN-2003 (first entry)
DE Prokaryotic essential gene #28741.
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX Staphylococcus epidermidis.
OS
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX
XX (BLIT-) ELITRA PHARM INC.

PI Doucette-Stamm L, Bush D;
XX WPI; 2004-580138/56.
DR P-PSDB; ADS07057.
XX
PT New isolated polypeptide and encoding nucleic acid derived from
PT Staphylococcus epidermidis, useful for diagnosing, preventing and/or
PT treating an S. epidermidis bacterial infection.
XX
PS Claim 5; SEQ ID NO 2580; 741pp; English.
XX
XX The invention describes an isolated nucleic acid comprising a nucleotide
CC sequence with any of 3772 fully defined nucleotide sequences (SEQ ID NO:
CC 1-3772) and encoding an Staphylococcus epidermidis polypeptide with any
CC of 3772 fully defined amino acid sequences (SEQ ID NO: 3772-7544) as
CC given in the specification. Also described are: a recombinant expression
CC vector; a cell comprising a recombinant expression vector of (1);
CC producing an S. epidermidis polypeptide; an isolated nucleic acid
CC comprising a nucleotide sequence of at least 8 nucleotides in length; a
CC vaccine composition for prevention or treatment of an S. epidermidis
CC infection, comprising a nucleic acid cited above and a carrier; treating
CC a subject for S. epidermidis infection; a recombinant or substantially
CC pure preparation of an S. epidermidis polypeptide or its fragment; a
CC vaccine composition for prevention or treatment of an S. epidermidis
CC infection; detecting the presence of a Staphylococcus nucleic acid in a
CC sample; a computer readable medium having recorded in it the nucleotide
CC sequences with SEQ ID NO: 1-3772 or its fragments; a computer based
CC system for identifying fragments of the Staphylococcus genome of
CC commercial importance; a computer based system for identifying fragments
CC of the Staphylococcus plasmids of commercial importance; identifying
CC commercially important nucleic acid fragments of the Staphylococcus
CC genome and/or plasmids; and identifying an expression modulating fragment
CC of the Staphylococcus genome and/or plasmids. The methods and
CC compositions of the present invention are useful for the diagnosis,
CC prevention and/or treatment of an Staphylococcus epidermidis bacterial
CC infection. This sequence encodes a S. epidermis protein of the invention.
XX
SQ Sequence 1008 BP; 404 A; 129 C; 188 G; 287 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5, 07e-154 Length: 1008
Score: 1690.00 Matches: 335
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-724-972A-6352 (1-335) x ADS03285 (1-1008)

QY 1 GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeu 20
Db 1 CGAGTGAATCAGTCAGAGGTTTAAAAATTTAAAGTGTAAATGGCTTATGTGTTTAA 60

QY 21 IleAlaThrAlaAlaCysGlyAsnAsnSerSerAsnSerSerLysGluSerLys 40
Db 61 ATTGCAACTGCAGCATGTGGAATAATAGTTCAGTAACTCAAGTAAGAGTCATCAAAA 120

QY 41 AspGlyValGluLysHisGluGluGlyThrLysValProLysHisProLysArg 60
Db 121 GATGGAGTGTGAATCAAGCACGAGAGAGGTACTACGAAAGTACTTAAACACCTAAACGT 180

QY 61 ValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProVal 80
Db 181 GTTGTGTTCTTCTGAGTATTCATTTGTGTCGCTTGTAGTGTGTTAAACCTGTT 240

QY 81 GlyIleAlaAspAsnLysLysAsnArgGlyIleLysProLeuArgLysIleGly 100
Db 241 GGGATAGCGGATGATACAAAAAAATCGTATTATTAAACCATTAAGAGTAATAATTGGA 300

QY 101 LysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluLysSerLysLeuLys 120
Db 301 AAATACACTTCTGTAGGAACACGTAAACCTAACTTAGAAGAAATCAGTAAACTTAAA 360

1000

319 AATTATCTCTCTAGGTACACGTAAACAGCCAACTTAGAGAAATTAGTAATTTAAAA 378
121 ProAspLeuIleIleAlaAspAenArgHisLysGlyIleTyrLysAspLeuAenLys 140
379 CCGGATTTAATATCGCTAGACAGTAAAGGTATTAAATAAGAAATTAACAAA 438
141 IleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAenIleAspAla 160
439 ATTGCACCAACATTATCATTAAGAGTTTGTAGGAGACTACAAACAAACATTAATTCG 498
161 PheLysThrIleSerLysAlaLeuLysGluGluGluLysLysArgLeuGluGlu 180
499 TTCAAAACAATTGCTAAAGCTTTAAATAAGAAAGAGCGGAAAGCGCTCTTGTCTGAA 558
181 HisAspLysLysIleGluGluTyrLysLysGluIleThrMetAspLysAenGlnLysVal 200
559 CATGATAAATTAATCAAAAAGTATAAAGATGAATTAAGTTTGTATAGAAATCAAAAAGTG 618
201 LeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAenSerTyrValGly 220
619 CTTCACGAGTTGTTGCTAAAGCTTTTATTAGCACATCCAACTTTCATATGTTGGA 678
221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu 240
679 CAATTTTAAACGAACTTGGATTAAAAATGCATTAACTGATGATGTAACAAAAGTTTA 738
241 SerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsn 260
739 AGTAAATCTTGAAGACCTTACTTCAATAGTACTGAACATTTAGCTGACTTAAT 798
261 ProGluArgMetPheIleMetThrAsnLysAlaSerSerAenGluProSerLeuLysGlu 280
799 CCTCAACGCTGATTAATTATGACAGATAATGCTAAAAAGATTTCTGCTGAATTCAGAGAG 858
281 LeuGluLysAspProValTyrLysLysLeuAsnAlaValLysAenGlnArgValAspIle 300
859 TTACAAGAGATCCAACTTGGAAAAAGTTGAACGAGTTAAAAATTAATCGCGTGGATTT 918
301 LeuAspArgAspLeuTyrAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLys 320
919 GTTGACCGTGATGTTTGGCAAGATCTCGTGCCTTAATTTCTTCTGAAGAAATGCTAAA 978
321 GluLeuValGluLeuSerLysLysAspSerLys 331
979 GAACCTTGTGTAATTATCAAAAAAGAGACAAAAG 1011
RESULT 4
ACC48532/c
ID ACC48532 standard; DNA; 2981 BP.
XX ACC48532;
XX AC
XX 11-AUG-2003 (first entry)
XX DT
XX DE Staphylococcal surface-exposed immunogenic polypeptide DNA.
XX KW Surface-exposed immunogenic polypeptide; SEIP; siderophore; receptor;
XX KW antibacterial; vaccine; gene; ds.
XX OS Staphylococcus aureus.
XX FH Location/Qualifiers
XX FT complement (968. .1951)
XX FT /*tag= a
XX FT /product= "SEIP"
XX PN W02003020875-A2.
XX PD 13-MAR-2003.
XX XX 17-JUN-2002; 2002WO-US019224.
XX PF 17-JUN-2001; 2001US-0298975P.
XX PR

23-MAY-2000; 2000US-0206848P.
26-MAY-2000; 2000US-0207272P.
23-OCT-2000; 2000US-0242578P.
27-NOV-2000; 2000US-0253625P.
22-DEC-2000; 2000US-0257931P.
16-FEB-2001; 2001US-0269308P.
(BLIT-) ELITRA PHARM INC.
Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
Yamamoto RT, Xu HH;
WPI; 2001-611495/70.
P-PSDB; AAU36660.
New polynucleotides for the identification and development of
antibiotics, comprise sequences of antisense nucleic acids.
Claim 27; SEQ ID NO 8156; 511pp; English.
The invention relates to antisense inhibitors of genes essential to
prokaryotic cellular proliferation, their use in identifying the genes,
their use in the discovery of novel antibiotics, the essential genes
themselves and the encoded proteins. The prokaryotes used are Escherichia
coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
useful for the identification of potential new targets for antibiotic
development. The antisense nucleic acids can also be used to identify
proteins used in proliferation, to express these proteins, and to obtain
antibodies capable of binding to the expressed proteins. The proteins can
be used to screen compounds in rational drug discovery programmes. The
antisense nucleic acid sequence is also useful to screen for homologous
nucleic acids which are required for cell proliferation in a wide variety
of organisms. The present sequence encodes an essential prokaryotic
cellular proliferation protein. Note: The sequence data for this patent
did not form part of the printed specification, but was obtained in
electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences
Sequence 1014 BP; 410 A; 126 C; 187 G; 291 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,18e-110 Length: 1014
Score: 1242.00 Matches: 239
Percent Similarity: 84.8% Conservatives: 42
Best Local Similarity: 72.21% Mismatches: 50
Query Match: 73.4% Indels: 1
DB: 4 Gaps: 0
US-10-724-972A-6352 (1-335) x AAS54519 (1-1014)
QY 1 GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeu 20
Db 20 GGAGTGGAAACG-ATGAGAGGCTTAAACACTTTTAGTATATGGGATTATAGTTGCTTA 78
QY 21 IleAlaThrAlaAlaCysGlyAenAenSerSerSerSerSerSerLysLysSerLys 40
Db 79 TTTTATAGTTGAGCTTGTGTAATACGATAATTTCAAGTAAAAAGAAATCATCAACATAA 138
QY 41 AspGlyValGluLysHisGluGluGlyThrThrLysValProLysHisProLysArg 60
Db 139 GATACATTATTTCGGTAAAGATGAAGATGGTACAGTAAAGTACCTAAAGATGCAAAACGT 198
QY 61 ValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProVal 80
Db 199 ATCGTTGATTAGTACTCTATTTCGAGATGCATTAGCAGCATTTAGACGTTAAACAGTT 258
QY 81 GlyIleAlaAspAspAsnLysLysAenArgIleIleLysProLeuArgAspLysIleGly 100
Db 259 GGTATTGCTGATCATGCTAAGAAAAACGATATCATTAACACAGTTAGAGAAAAATGGG 318
QY 101 LysTyrThrSerValGlyThrArgLysGlnProAenLeuGluLysLeuLys 120

XX (DSQU-) D-SQUARED BIOTECHNOLOGIES INC.
XX Scott DL;
XX WPI; 2003-300870/29.
XX P-PSDB; ABR41865.
XX Novel surface-exposed immunogenic polypeptide of Staphylococcus aureus
XX containing receptors for siderophores or iron-binding ligands, useful for
XX producing antibodies effective against Staphylococci infection.
XX Claim 4; Page 53; 62pp; English.
XX
XX The present sequence is that of DNA encoding novel Staphylococcus aureus
XX surface-exposed immunogenic polypeptide (SEIP) D2 SA03. To isolate SEIP
XX genes, an expression library of S. aureus genomic DNA was screened with
XX anti-SEIP antisera. DNA was isolated from clones expressing SEIPs, and
XX amino acid sequences were deduced and analysed for conserved and/or
XX functional domains. The SEIP encoded by the present DNA sequence has
XX sequence homology with the siderophore family of periplasmic binding
XX proteins. The invention provides methods for the identification,
XX production and recovery of SEIPs. The SEIPs can be used individually, or
XX in combination, to produce anti-staphylococcal antibodies useful in
XX passive or active immunisation strategies to prevent or contain
XX staphylococcal infection. They can also be used to develop diagnostic
XX assays
XX
XX Sequence 2981 BP; 988 A; 559 C; 386 G; 1048 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3,36e-109 Length: 2981
Score: 1233.00 Matches: 238
Percent Similarity: 84.59% Conservative: 42
Best Local Similarity: 71.90% Mismatches: 51
Query Match: 72.96% Indels: 1
DB: 8 Gaps: 0

US-10-724-972A-6352 (1-335) x ACC48532 (1-2981)

QY 1 GluValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeu 20
DB 1992 GGAGTGGAAACG-ATGAGAGGTCTAAACACTTTTAGTATATATGGGATTATGCTTGA 1934
QY 21 IleAlaThrAlaAlaCysGlyAenAenSerSerAenSerSerLysGluSerLys 40
DB 1933 CTTTATAGTTCGAGTTCGTGTAATACGGATAATTCAGTAAAGAAAAGATCATCACTAAA 1874
QY 41 AspGlyValGluIleLysHisGluGluGlyThrThrLysValProLysHisProLysArg 60
DB 1873 GATACATATTCGTGTAATACGGATAATTCAGTAAAGAAAAGATCATCACTAAA 1814
QY 61 ValValValLeuGluThrSerPheValAlaLeuValAlaLeuAspValLysProVal 80
DB 1813 ATCGTTGTATATAGTACTCATTTGCAGATGCATTAGCAGCATTAAGCGTTAAACAGTT 1754
QY 81 GlyIleAlaAspAenLysLysAenArgIleLysProLeuArgAspLysIleGly 100
DB 1753 GGTATTCGTGATGATGTTAAGAAAACGATCATTAACACCGATTAGAGAAAATTCGG 1694
QY 101 LysThrThrSerValGlyThrArgLysGlnProAenLeuGluGluIleSerLysLys 120
DB 1693 GATTATACCTCTGTAGGTACAGTAAACAGCCAACTTAGAGGAATTAATTAATAA 1634
QY 121 ProAspLeuIleIleAlaAspAenArgHisLysGlyIleTyrLysAspLeuAenLys 140
DB 1633 CCGGATTTAATATTCGCTGATAGATAGACATAAAGGTATTAATAAGAAATTAACAAA 1574
QY 141 IleAlaProThrIleGluLeuLysSerPheAspGlyAspThrAenGluAenIleAspAla 160
DB 1573 ATTCACCAACATATATCAATTAAGAGTTTGTATGAGACTACAAACAAAATTAATTCG 1514
QY 161 PheLysThrIleSerLysAlaLeuGlyLysGluGluGluGlyLysArgLeuGluGlu 180

DB 1513 TTCAACAATTCGTAAGCTTTAAATAAGAAAAGAACGCGAAAACGCTTTCCTGAG 1454
QY 181 HisAspLysLysIleGluGluTyrLysLysGluIleThrMetAspLysAenGlnLysVal 200
DB 1453 CATGATAAATTAATCAATAAGTATAAAGATGAATTAATTAATTTGATAGAAATCAAAAGTG 1394
QY 201 LeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAenSerIyrValGly 220
DB 1393 CTTCCAGCAGTAGTTCGTAAGCTGTTTATTAGCACATCCAACTATTCATATGTTGA 1334
QY 221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu 240
DB 1333 CAATTTTAAACGAACCTAGGATTTAAATGCAATTAAGTCAGATGAACAAAGGTTTA 1274
QY 241 SerLysTyrIleLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsn 260
DB 1273 AGTAAATATTTGAAGGACCTTACTTACAATTAGACACTGAACTTAGCTGATTAAAT 1214
QY 261 ProGluArgMetPheIleMetThrAsnLysAlaSerSerAenGluProSerLeuLysGlu 280
DB 1213 CCAGACGATGATGATCATTTATGACAGATCATGTAAAAAAGATTCGCTGAATTCAGAAG 1154
QY 281 LeuGluLysAspProValTyrLysLysLeuAenAlaValLysAenGlnArgValAspIle 300
DB 1153 TTACAGAGATGCAACATGGAAGAAAGTTGAATGCAAGTTAAATAATTCGCTGATATT 1094
QY 301 LeuAspArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLys 320
DB 1093 GTTGACCGTATGTTGGGCAAGATCTCGTGGCTTAATTTCTCTGAAGAAATGGCTAAA 1034
QY 321 GluLeuValGluLeuSerLysLysAspSerLys 331
DB 1033 GAACTTGTGTAATTTATCAAAAAAAGAACAAAG 1001

RESULT 5

ACA19774

ID ACA19774 standard; DNA; 984 BP.

XX ACA19774;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #1431.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.

XX Staphylococcus aureus.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US0009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;XX WPI; 2003-029926/02.
XX P-PSDB; ABU15904.XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.

PT preventing Staphylococcal infection, specifically an infection caused by
 XX *S. aureus*, e.g. *sepsis*.

PS Claim 6; SEQ ID NO 5331; 49pp; English.

XX... The invention relates to novel genes and encoded proteins from
 CC *Staphylococcus aureus*. A composition comprising the *S. aureus* protein, a
 CC nucleic acid encoding the protein, or an antibody to the protein, is
 CC useful as a pharmaceutical, particularly as a vaccine for treating or
 CC preventing infection due to *Staphylococcus aureus*, specifically an
 CC infection caused by *S. aureus*. The composition is particularly useful for
 CC treating or preventing sepsis in a patient. The composition can also be
 CC used for diagnostics. The protein is also used in an assay for enzymatic
 CC studies and as a target for antibiotics. This sequence represents one of
 CC the novel *S. aureus* genes of the invention

SQ Sequence 1044 BP; 425 A; 128 C; 191 G; 300 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 6.25e-108 Length: 1044
 Score: 1214.00 Matches: 237
 Percent Similarity: 84.59% Conservative: 43
 Best Local Similarity: 71.60% Mismatches: 51
 Query Match: 71.83% Indels: 2
 DB: 8 Gaps: 0

US-10-724-972a-6352 (1-335) x ACF74986 (1-1044)

QY 1 GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuPheValLeu 20
 DB 54 CGAGTGGAAAG-ATGAGAGGTCTAAACCTTT-AGTATATTGGATTAATGTCCTTA 111
 QY 21 IleAlaThrAlaAlaCysGlyAenAenSerSerSerSerSerSerSerSerSerSerLys 40
 DB 112 CTTTATTGTCAGCTTGTGTAATACGGTAATTCAGTAATAAAGAAAGAAATCATCACTAAA 171
 QY 41 AspGlyValGluIleLysHisGluGluGlyThrLysValProLysHisProLysArg 60
 DB 172 GATACATTTTCGGTAAAGATGAAATGGTACAGTAAAGTACCTAAAGATCAAAACGT 231
 QY 61 ValValValLeuGluTyrSerPheValAlaLeuValAlaLeuValAlaLeuValAlaLeu 80
 DB 232 ATCGTTGTTAGTAGTACTCATTTGCAGATGCTAGCAGCAATTAAGCTTAAACCCAGTT 291
 QY 81 GlyIleAlaAspAspAspLysLysLysLysLysLysLysLysLysLysLysLysLysLys 100
 DB 292 GGTATTGCTGATGATGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 351
 QY 101 LysTyrThrSerValGlyThrArgLysGlnProAenLeuGluGluLysLysLysLys 120
 DB 352 GATTATCTTCTGTAGGTACACGTAAACAGCCAACTTAGAAGAAATTAGTAAATTTAAA 411
 QY 121 ProAspLeuIleIleAlaAspAenAenArgHisLysGlyIleTyrLysAspLeuAenLys 140
 DB 412 CCGGATTTAATATCGCTGATAGCAGTACATAGAGTATTAATAAGAAATTAACAAA 471
 QY 141 IleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAenGluAenIleAspAla 160
 DB 472 ATTGCACCAATATCATCAATAGAGTTTTCATGGAGACTCAACAAATTAATTAATTCG 531
 QY 161 PheLysThrIleSerLysAlaLeuGlyLysGluGluGlyLysLysLysLysLysLysLys 180
 DB 532 TTCAAAAACAAATTGCTAAAGCTTTAAATAAGAAAGAAAGAAAGAAAGAAAGCTTCTG 591
 QY 181 HisAspLysLysIleGluGluTyrLysLysGluIleThrMetAspLysAenGlnLysVal 200
 DB 592 CATGATAAATTAATCAATAGATTAAGATTAATTAATTTTGTAGAAATCAAAAAGTG 651
 QY 201 LeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAenSerTyrValGly 220
 DB 652 CTTCCAGCAGTAGTTGCTAAAGCTGGTTTATTAGCATCAATCAATCAATCAATCAATCA 711
 QY 221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu 240

DB 712 CAATTTTAAACGAACTAGGATTTAAATAATGCATTAAAGTACGATGTAACAAAAGGTTTA 771
 QY 241 SerLysTyrLeuLysGlyProTyrLeuGlnMetAenThrGluThrLeuSerGlnValAen 260
 DB 772 AGTAAATATTTGAAAGGACCTTACTTACAATTAGACACTGAACATTTAGCTGATTAAAT 831
 QY 261 ProGluArgMetPheIleMetThrAenLysAlaSerSerAenGluProSerLeuLysGlu 280
 DB 832 CCAGAGCGTATGATCATTTATGACAGATCATGCTAAATAAGAGATTCTGCTGAATTCAGAAG 891
 QY 281 LeuGluLysAspProValTyrLysLysLeuAenAlaValLysAenGlnArgValAspIle 300
 DB 892 TTCAAGAAGATGCAACATGCAAAAGTTGAATGCAGTTAAATAATATCGGTGGATATT 951
 QY 301 LeuAspArgAspLeuTrpAlaArgSerArgGlyLeuLysSerGluGluMetAlaLys 320
 DB 952 GTTGACCGTATGTTGGGCAAGATCTCGTGGCTTAATTTCTTCTGAAGAATGGCTAAA 1011
 QY 321 GluLeuValGluLeuSerLysLysAspSerLys 331
 DB 1012 GAACCTGTTGAATTTATCAAAAAAAGAACAAAAG 1044

RESULT 7

AAS51822

ID AAS51822 standard; DNA; 927 BP.

AC AAS51822;

DT 13-FEB-2002 (first entry)

XX Staphylococcus aureus DNA for cellular proliferation protein #239.
 DE Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
 KW antibacterial; drug design.

XX Staphylococcus aureus.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US0009180.

XX 21-MAR-2000; 2000US-0191078P.

XX 23-MAY-2000; 2000US-0206848P.

XX 26-MAY-2000; 2000US-0207727P.

XX 23-OCT-2000; 2000US-0242578P.

XX 27-NOV-2000; 2000US-0253625P.

XX 22-DEC-2000; 2000US-0257931P.

XX 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 XX Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

XX P-PSDB; AAU33963.

XX New polynucleotides for the identification and development of
 XX antibiotics, comprise sequences of antisense nucleic acids.

XX Claim 27; SEQ ID NO 4404; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
 CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
 CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify

CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence encodes an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 927 BP; 370 A; 120 C; 169 G; 268 T; 0 U; 0 Other;

Alignment Scores: 8.42e-105 Length: 927
Pred. No.: 1181.00 Matches: 225
Score: 85.11% Conservative: 38
Percent Similarity: 85.11% Mismatches: 46
Best Local Similarity: 72.82% Indels: 0
Query Match: 69.88% Gaps: 0
DB: 4

US-10-724-972A-6352 (1-335) x AAS51822 (1-927)

Qy 16 LeuLeuPheValLeuLeuLeuAlaAlaCysGlyAenAenSerSerAenSerSer 35
Db 1 TTAATAGTTCCTTATTTAGTTGCGAGCTTGTGTAATACGGATAATTCAGTAAAAA 60
Qy 36 LysGluSerSerLysAspGlyValGluLeuLysHisGluGluGlyThrLysValPro 55
Db 61 GAATCATCAACTAAGATACCTATTTCGGTAAAGATGAAATGCTACAGTAAAGTACCT 120
Qy 56 LysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeu 75
Db 121 AAAGATGCAAAACGATCGTTGTTATTAGTACTCATTTGCAGATGCGATTACGACATTA 180
Qy 76 AspValLysProValGlyLeuAlaAspAspAsnLysLysAsnArgGileileLysProLeu 95
Db 181 GACGTTAAACAGTGGTATTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 240
Qy 96 ArgAspLysLysGlyLysThrSerValGlyThrArgLysGlnProAsnLeuGluGlu 115
Db 241 AGAGAAAAATTTGGGAATTTATCTTCTGTAGGTACACGTTAAACACCAACTTAGAGGAA 300
Qy 116 IleSerLysLeuLysProAspLeuLeuLeuAlaAspAsnAsnArgHisLysGlyLys 135
Db 301 ATTAGTAATTAACCCGATTTAATTCGTCGATGATGATGATGATGATGATGATGATGAT 360
Qy 136 LysAspLeuAsnLysLysLeuAlaProThrLysGluLysSerPheAspGlyAspTyrAen 155
Db 361 AAAGAATTAACAAAAATTCACCAACATTTATCATTAAGAGAGTTTGTATGGAGACTACAA 420
Qy 156 GluAsnLysAspAlaPheLysThrLysSerLysAlaLeuGlyLysGluGluGlyLys 175
Db 421 CAAAACATTAATTCGTTCAAAACAATTCGTTAAAGCTTTAAATTAAGAAAAAGACGAA 480
Qy 176 LysArgLeuGluGluHisAspLysLysLysGluGluTyrLysLysGluLeuThrMetAsp 195
Db 481 AAGCGTCTTGCTGAACATGATTAATTAATCAAAAGTATAAAGATCAAAATTAAGTTGAT 540
Qy 196 LysAsnGlnLysValLeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHisProSer 215
Db 541 AGAAATCAAAAAGTGTCTCCAGCAGTTGTTGCTTAAAGCTGTTTATAGCACATCCAAAC 600
Qy 216 AsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAsp 235
Db 601 TATTTCATGTTGGACAAATTTTAAACGAACCTTGGATTTAAATGCAATTAAGTATGAT 660
Qy 236 ValThrLysGlyLeuSerLysThrLysGlyProTyrLeuGlnMetAsnThrGluThr 255
Db 661 GTACAAAAGGTTTAACTAATCTTCAAGAGGACCTTACTTACAATTAGATACCTGACAT 720
Qy 256 LeuSerGlnValAenProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGlu 275

Db 721 TTAGCTGACTTAATTCCTGAACGCAATGATTATTATGACAGATAATGCTAAAAAGATTCT 780
Qy 276 ProSerLeuLysGluLeuGluLysAspProValTrrPlyLysLysLeuAsnAlaValLysAen 295
Db 781 GCTGATTCAGAGAGTTTACAGAGATCCAACTTGGAAAAAGTTGAAACGCACTTAAAAAT 840
Qy 296 GlnArgValAspLysLeuAspArgAspLeuTrrAlaArgSerArgGlyLeuLysSerSer 315
Db 841 AATCGCGTGGATATTGTTGACCGTGATGTTGGCAAGATCTCGTGGCTTAATTTCTTCT 900
Qy 316 GluGluMetAlaLysGluLeuValGlu 324
Db 901 GAAGAAATGCTAAAGAACTTGTGAA 927

RESULT 8

AAS51612 ID AAS51612 standard; DNA; 957 BP.

XX AAS51612; AC

XX 13-FEB-2002 (first entry)

XX Staphylococcus aureus DNA for cellular proliferation protein #29.

XX Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
antibacterial; drug design.

XX Staphylococcus aureus.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.

XX 21-MAR-2000; 2000US-0191078P.

XX 23-MAY-2000; 2000US-0206848P.

XX 26-MAY-2000; 2000US-0207727P.

XX 23-OCT-2000; 2000US-0242578P.

XX 27-NOV-2000; 2000US-0253625P.

XX 22-DEC-2000; 2000US-0257931P.

XX 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
WPI; 2001-611495/70.
P-PSDB; AAU33753.

XX New polynucleotides for the identification and development of
antibiotics, comprise sequences of antisense nucleic acids.

XX Claim 27; SEQ ID NO 4194; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
prokaryotic cellular proliferation, their use in identifying the genes,
their use in the discovery of novel antibiotics, the essential genes
themselves and the encoded proteins. The prokaryotes used are Escherichia
coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
useful for the identification of potential new targets for antibiotic
development. The antisense nucleic acids can also be used to identify
proteins used in proliferation, to express these proteins, and to obtain
antibodies capable of binding to the expressed proteins. The proteins can
be used to screen compounds in rational drug discovery programmes. The
antisense nucleic acid sequence is also useful to screen for homologous
nucleic acids which are required for cell proliferation in a wide variety
of organisms. The present sequence encodes an essential prokaryotic
cellular proliferation protein. Note: The sequence data for this patent
did not form part of the printed specification, but was obtained in
electronic format directly from WIPO at

Pred. No.:		9,63e-61	Length:	960
Score:		726.00	Matches:	153
Percent Similarity:		65.85%	Conservative:	61
Best Local Similarity:		47.08%	Mismatches:	99
Query Match:		42.96%	Indels:	12
DB:		4	Gaps:	5
US-10-724-972A-6352 (1-335) x AAS54387 (1-960)				
QY	6	ArgGlyLeuLysIleLeuSerValIleGlyLeuPheValLeuIleAlaThrAlaAla	25	
DB	7	AGGAATATCGTTAAATAGTTGTTATGCTTAATCTTCGTT---GTAGCAGTAGCGGT	63	
QY	26	CysGlyAsnAsnSerSerAsnSerSerLysGlySerLysAspGlyValGluIle	45	
DB	64	TGTGGTCAAAAGACTACT-----GAAGAGAAAACCTGAATGACGACAATA	108	
QY	46	LysHieGluGluGlyThrThrLysValProLysHieProLysArgValValLeuGlu	65	
DB	109	AAAGATGAATTAGGAACCTGAAAATAAAGAAAATCCTAAACGTGTTGTTGTTATTAGAA	168	
QY	66	TyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsp	85	
DB	169	TATAGTTTTCGTGATTATTAGCAGCATATAGATATGAACCTGTTGGTATTGCAGATGAT	228	
QY	86	AsnLysLysAsnArgIleIleLysProLysAspLysIleGlyLysTyrThrSerVal	105	
DB	229	GGCAGCAGTAAATAATACAAAGTCAGTAAGAGATAAGGTGGCGCATATGAATCGGTT	288	
QY	106	GlyThrArgLysGlnProAsnLeuGluGluIleSerLysLysProAspLeuIleIle	125	
DB	289	GGATCTAGCGCAACGCAATATGGAAGTGATAAGTAAATTAACCGGATTTGATCATTA	348	
QY	126	AlaAspAsnAsnArgHieLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIle	145	
DB	349	GCAGATGTTAGCAGACATAGAAAATCAATCAGAAATTAAGCAAAATTTGCTCCGACAATC	408	
QY	146	GluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSer	165	
DB	409	ATGTTAGTACGGGTACGGGAGATTATAATGCAAAATATTGATGCAATTTAAACACAGTCGCT	468	
QY	166	LysAlaLeuGlyLysGluGluGluGlyLysLysArgLeuGluGluHieAspLysLysIle	185	
DB	469	AAAGCAGTAGCAAGAGAGAAGAGAGCGGAGAAACGCTCGAAAGCATGATAAATATTA	528	
QY	186	GluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAla	205	
DB	529	CGCGAGATTAGAAAGAAATTTGAACAGAGTAGCTTAAACCTGCATTTGCGATTCGATC	588	
QY	206	AlaLysSerGlyLeuLeuAlaHieSerSerAsnSerTyrValGlyGlnPheLeuSerGln	225	
DB	589	TCAAGACAGGTATGTTTATTATTAATGAAGATACATTTATGGGACAAATTTCTTAATTAA	648	
QY	226	LeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLysSerLysTyrLeuLys	245	
DB	649	ATGGGTATTCACCTCGAAGTCAAAAGACAAAACCTACGCATGTTGGTGAAGCGCAAGGT	708	
QY	246	GlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPhe	265	
DB	709	GCTCTCTATTATTATTATTAATTAATGAGAACTTGCCATATCAATCCAAAGTTATGATT	768	
QY	266	IleMetThrAsn---LysAlaSerSerAsnGluProSerLeuLysGluLeuGluLysAsp	284	
DB	769	TTAGCCACTGACGGAAAAACGGCAAAAATAATAGAACGAAATTC-----ATTGAT	816	
QY	285	Pro---ValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg	303	
DB	817	CCTGCAGTTTGGAAATTCATTAAGAGCTGTGAAAGATATAACAAAGTTTATGACGTTGACCGA	876	
QY	304	AspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuVal	323	
DB	877	AATAAGTGGTTGAATCAAGGGGTATATTCGCAAGTGAAAGTATGCGCAGAAAGATTAGAA	936	

QY

324 GluLeuSerLysLys 328

DB

937 AAAATTGCAGAAAA 951

RESULT 10

ABT15015

ID

ABT15015 standard; DNA; 957 BP.

XX

ABT15015;

XX

06-MAR-2003 (first entry)

XX

Pathogen specific antigen related staphylococcal DNA SEQ ID No 301.

DE

Antibacterial; virucide; fungicide; protozoacide; cytostatic; anti-HIV;

KW

hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;

KW

auto-immunity; vaccine; staphylococcal infection; antibody; cancer;

KW

autoimmune disease; HIV; hepatitis; gene; ds.

XX

Staphylococcus sp.

OS

W0200259148-A2.

XX

01-AUG-2002.

XX

21-JAN-2002; 2002WO-EP000546.

XX

26-JAN-2001; 2001AT-00000130.

XX

(CIST-) CISTEM BIOTECHNOLOGIES GMBH.

PA

Meinke A, Nagy E, Von Ahnen U, Klade C, Henics T, Zauner W;

XX

Minh DB, Vytvytska O, Etz H, Dryla A, Weichhart T, Hafner M;

PI

Tempelmaier B;

PI

WPI; 2003-075410/07.

DR

Identifying, isolating and producing hyperimmune serum-reactive antigens

XX

from a pathogen, for preparing vaccine or medicament for treating or

PT

preventing e.g. staphylococcal infections, comprises providing antibody

PT

preparation.

XX

Example 7; Page 203; 252pp; English.

XX

The invention relates to a novel method for identifying, isolating and

CC

producing hyperimmune serum-reactive antigens from a pathogen, tumour,

CC

allergen, a tissue or host prone to auto-immunity, where the antigens are

CC

used in a vaccine, comprises providing antibody preparation from a plasma

CC

pool of a type of animal, or individual sera with antibodies against the

CC

specific pathogen, tumour, allergen, tissue or host prone to auto-

CC

immunity. The hyperimmune serum-reactive antigens comprising any of the

CC

62 sequences of 53-2261 amino acids fully defined in the specification,

CC

or their hyperimmune fragments are useful for the manufacture of a

CC

pharmaceutical preparation, particularly a vaccine against staphylococcal

CC

infections or colonisation against S. aureus or S. epidermidis. The

CC

preparation of antibodies is useful for the manufacture of a medicament

CC

for treating or preventing staphylococcal infections or colonisation

CC

against S. aureus or S. epidermidis. The antibody preparations may also

CC

be used for diagnostic and imaging purposes. Other conditions that can be

CC

treated include cancer, autoimmune diseases or infections caused by viral

CC

(e.g. HIV, hepatitis A, B or C), fungal or protozoan pathogens. This

CC

polynucleotide sequence represents staphylococcal DNA relating to the

CC

method for identifying and producing pathogen specific antigens of the

CC

invention

XX

Sequence 957 BP; 383 A; 116 C; 210 G; 248 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

Length:

Matches:

Conservative:

Mismatches:

Indels:

1,34e-60

724.50

66.67%

46.86%

42.87%

957

149

63

95

11

DB: 8 Gaps: 4

US-10-724-972A-6352 (1-335) x ABT15015 (1-957)

Qy 13 ValletGlyLeuLeuPheValLeuLeuAlaAlaCysGlyAsnAsnSerSerSer 32
 Db 25 GTTGTGTTGCTAAATTTAGTTGTAGCAGTGGGTTGCTGCTCAAAAGATACT--- 81
 Qy 33 AsnSerSerLysGluSerSerLysAspGlyValGluLeuLysHisGluGluGlyThr 52
 Db 82 -----GAAGAGAAACCTGAATGACGCAATAAAGATGAATAGGAACCTCAA 129
 Qy 53 LysValProLysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeu 72
 Db 130 AAAATTAAAGAAATCTTAAACGTTGTTGTTAGTAATATAGTTTGTCTGATTATTA 189
 Qy 73 ValAlaLeuAspValLysProValGlyLeuAlaAspAsnLysLysAsnArgIleile 92
 Db 190 GCAGCATTAGATGAACCTGTTGTTATTCAGATGTCGACACTAAATATATACA 249
 Qy 93 LysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsn 112
 Db 250 AATGATGATAGATAGATAGATGCGGCATATCAATCGGTTGATCTAGACCGCAACGAAT 309
 Qy 113 LeuGluGluLysSerLysLeuLysProAspLeuLeuLeuLeuLeuLeuLeuLeuLys 132
 Db 310 ATCGAAGTATAGTAAATTAACACCGGATTTGATCATTCAGATGTTAGCAGACATAAG 369
 Qy 133 GlyIleTyrLysAspLeuAnLysIleAlaProThrIleGluLeuLysSerPheAspGly 152
 Db 370 AAAATCAATCAAGATTGACAAATTTGCTCCGACCAATCATGTTAGTCGGTACGGGA 429
 Qy 153 AspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 172
 Db 430 GATTATAATGCAAAATTAATGAAGCATTTAAACAGTCGCTAAAGCAGTAGCGCAAGAGAAA 489
 Qy 173 GluGlyLysLeuArgLeuGluGluHisAspLysLysLysLysLysLysLysLysLys 192
 Db 490 GAAGCGCAGACGCTCTGAAAGACATGATAAATATTAGCGGAGATTAGAAAGAAATTT 549
 Qy 193 ThrMetAspLysAsnGlnLysValLeuProAlaValAlaAlaLysSerGlyLeuLeuAla 212
 Db 550 GAACAGAGTACGTTAAATCTGCATTTGCTTCGATCTCAAGACGAGTATGTTTATT 609
 Qy 213 HisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232
 Db 610 AATAATGAGATACATTTATGGGACAAATCTTAATTAATTAATTAATTAATTAATTAAT 669
 Qy 233 SerAspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsn 252
 Db 670 ACAAAAGACAAACTACCGATGTTGGTGAACGCAAGGGTGGCTCTATATATATTAAT 729
 Qy 253 ThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn---LysAla 271
 Db 730 AATGAAGAACTTGCCAAATATCAATCCAAAGTTATGATTTAGCCACTGACGGAACG 789
 Qy 272 SerSerAsnGluProSerLeuLysGluLeuGluLysAspPro---ValTyrLysLysLeu 290
 Db 790 GACAAAAATAGAACCAATTC-----ATTGATCTCGCAGTTTGGAAATCATTA 837
 Qy 291 AsnAlaValLysAsnGlnArgValAspLeuLeuAspArgAspLeuTyrPalaArgSerArg 310
 Db 838 AAAGCTGGAAGATTAACAAAGTTTATGACGCTTGACCGGAAATAGTGGTTGAAATCAAGG 897
 Qy 311 GlyLeuLysSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328
 Db 898 GGGATTATCGCAGTGAAGTATGCGCAGAGATTTAGAAAAAATTTGCGAGAAAAA 951

RESULT 11

ACA20079
 ID ACA20079 standard; DNA; 960 BP.
 XX
 AC ACA20079;

XX 19-JUN-2003 (first entry)
 XX Prokaryotic essential gene #1736.
 XX Antisense; db; prokaryotic essential gene; cell proliferation;
 XX drug design; gene.
 XX Staphylococcus aureus.
 XX WO200277183-A2.
 XX 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 XX 06-SEP-2001; 2001US-00948993.
 XX 25-OCT-2001; 2001US-0342923P.
 XX 08-FEB-2002; 2002US-00072851.
 XX 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI: 2003-029926/02.
 XX P-PSDB; ABU16209.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 XX for homologous nucleic acids required for cellular proliferation to
 XX isolate candidate molecules for rational drug discovery programs.
 XX Claim 14; SEQ ID NO 7949; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 XX the 6213 antisense sequences given in the specification where expression
 XX of the nucleic acid inhibits proliferation of a cell. Also included are:
 XX (1) a vector comprising a promoter operably linked to the nucleic acid
 XX encoding a polypeptide whose expression is inhibited by the antisense
 XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
 XX polypeptide or its fragment whose expression is inhibited by the
 XX antisense nucleic acid; (4) an antibody capable of specifically binding
 XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 XX proliferation or the activity of a gene in an operon required for
 XX proliferation; (7) identifying a compound that influences the activity of
 XX the gene product or that has an activity against a biological pathway
 XX required for proliferation, or that inhibits cellular proliferation; (8)
 XX identifying a gene required for cellular proliferation or the biological
 XX pathway in which a proliferation-required gene or its gene product lies
 XX or a gene on which the test compound that inhibits proliferation of an
 XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
 XX compound's activity; (11) a culture comprising strains in which the gene
 XX product is overexpressed or underexpressed; (12) determining the extent
 XX to which each of the strains is present in a culture or collection of
 XX strains; or (13) identifying the target of a compound that inhibits the
 XX proliferation of an organism. The antisense nucleic acids are useful for
 XX identifying proteins or screening for homologous nucleic acids required
 XX for cellular proliferation to isolate candidate molecules for rational
 XX drug discovery programs, or for screening homologous nucleic acids
 XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 XX prokaryotic essential genes. Note: The sequence data for this patent did
 XX not form part of the printed specification, but was obtained in
 XX electronic format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences

Sequence 960 BP; 385 A; 115 C; 210 G; 250 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.34e-60 Length: 960
 Score: 724.50 Matches: 149

Percent Similarity: 66.67% Conservative: 63
 Best Local Similarity: 46.86% Mismatches: 95
 Indels: 11
 Query Match: 42.87%
 DB: 4
 Gaps: 4

US-10-724-972A-6352 (1-335) x ACA20079 (1-960)

QY 13 ValIleGlyLeuPheValLeuIleAlaThrAlaAlaCysGlyAsnAsnSerSerSer 32
 Db 25 GTTGTGTTTATCTAATCTAGTTAGTACGAGTGGGTTGTGGTCAAAAAGATACT--- 81
 QY 33 AsnSerSerLysGluSerSerLysAspGlyValGluIleLysHisGluGluGlyThrThr 52
 Db 82 -----GAAGAGAAAACCTGAAATGACGACAAATAAAGATGAATTAGGAACGAA 129
 QY 53 LysValProLysHisProLysArgValValLeuGluTyrSerPheValAspAlaLeu 72
 Db 130 AAATTAAGAAAATCCTTAACGTGTGTGTATTAGAAATATAGTTTGTCTGATTATTTA 189
 QY 73 ValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLysLysAsnArgIleIle 92
 Db 190 GCAGCATTAGATGAACCTGTTGTTATTCAGATGATGGCAGCACTAAAAATATAACA 249
 QY 93 LysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsn 112
 Db 250 AAGTCAGTAAGACATGAATGGGCGATATGAATCGTTGGATCTAGACCGCAACCGAAT 309
 QY 113 LeuGluLysLysSerLysLeuLysProAspLeuIleAlaAspAsnAsnArgHisLys 132
 Db 310 ATGGAGTGAATAGTAAATTAACCGGATTTGATCATTCGACATGTTAGCAGACATAAG 369
 QY 133 GlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGly 152
 Db 370 AAAATCAATCAGATTAAGTAAATTTGCAAAATTTGCTCCGCAATCATGTTAGTCGGTGGCA 429
 QY 153 AspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 172
 Db 430 GATTATATGCAATATTGAAGCATTTTAAACAGTCGCTAAACAGTAGGCAAGAGAAA 489
 QY 173 GluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGluTyrLysLysGluIle 192
 Db 490 GAAGCGAGAGCGCTCTGGAAAGCATGATAAATATTACCGAGATTAAGAAAGAAAAT 549
 QY 193 ThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLysLeuAla 212
 Db 550 GAACAGAGTAGCTTAAAAATCTGCATTTGCATTCGGTATCTCAAGAGCAGGTATGTTATT 609
 QY 213 HisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232
 Db 610 AATAATGAAGATACATTTATCGGCAATCTTAATTAATGGGTATTCACCTGAAGTC 669
 QY 233 SerAspAspValThrLysGlyLysSerLysTyrLeuLysGlyProTyrLeuGlnMetAsn 252
 Db 670 ACAAAAGCAAACTACCGCATGTTGGTGAACGCAAGGCTGCTCTATATATATTTAAAT 729
 QY 253 ThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn---LysAla 271
 Db 730 AATGAAGAACTTGCCAATATCAATCCAAAGTTATGATTTTACGCCTGACGGAAGAACG 789
 QY 272 SerSerAsnGluProSerLeuLysGluLeuGluLysAspPro---ValTrpLysLysLeu 290
 Db 790 GACAAAATATAGAACGAATTC-----ATTGATCTCGAGTTTGGAAATCATTA 837
 QY 291 AsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaAspSerArg 310
 Db 838 AAGCTGTGAAAGATAACAAGTTTATGCGTTTGACCGCAATAAGTGGTTCAAAATCAAGG 897
 QY 311 GlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328
 Db 898 GGGATTATCGCAAGTGAAAGTATGGCAGAAAGATTAGAAAAAATTTGCAGAAAAA 951

ACF74415 standard; DNA; 975 BP.
 ACF74415;
 20-NOV-2003 (first entry)
 Staphylococcus aureus DNA #2095.
 Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
 enzymatic assay; antibiotic target; gene; ds.
 Staphylococcus aureus.
 WO200294868-A2.
 28-NOV-2002.
 27-MAR-2002; 2002WO-IB002637.
 27-MAR-2001; 2001GB-00007661.
 (CHIR-) CHIRON SPA.
 Massignani V, Mora M, Scarselli M;
 WPI: 2003-120786/11.
 P-PSDB; ABM72855.
 New Staphylococcus aureus protein, useful as a vaccine for treating or
 preventing Staphylococcal infection, specifically an infection caused by
 S. aureus, e.g. sepsis.
 Claim 6; SEQ ID NO 4189; 49pp; English.
 The invention relates to novel genes and encoded proteins from
 Staphylococcus aureus. A composition comprising the S. aureus protein, a
 nucleic acid encoding the protein, or an antibody to the protein, is
 useful as a pharmaceutical, particularly as a vaccine for treating or
 preventing infection due to Staphylococcus bacteria, specifically an
 infection caused by S. aureus. The composition is particularly useful for
 treating or preventing sepsis in a patient. The composition can also be
 used for diagnostics. The protein is also used in an assay for enzymatic
 studies and as a target for antibiotics. This sequence represents one of
 the novel S. aureus genes of the invention
 Sequence 975 BP; 387 A; 116 C; 216 G; 256 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.37e-60 Length: 975
 Score: 724.50 Matches: 149
 Percent Similarity: 66.67% Conservative: 63
 Best Local Similarity: 46.86% Mismatches: 95
 Query Match: 42.87% Indels: 11
 DB: 4 Gaps: 4

US-10-724-972A-6352 (1-335) x ACF74415 (1-975)

QY 13 ValIleGlyLeuPheValLeuIleAlaThrAlaAlaCysGlyAsnAsnSerSerSer 32
 Db 43 GTTGTGTTTATCTAATCTAGTTAGTACGAGTGGGTTGTGGTCAAAAAGATACT--- 99
 QY 33 AsnSerSerLysGluSerSerLysAspGlyValGluIleLysHisGluGluGlyThrThr 52
 Db 100 -----GAAGAGAAAACCTGAAATGACGACAAATAAAGATGAATTAGGAACGAA 147
 QY 53 LysValProLysHisProLysArgValValLeuGluTyrSerPheValAspAlaLeu 72
 Db 148 AAATTAAGAAAATCCTTAACGTGTGTGTATTAGAAATATAGTTTGTCTGATTATTTA 207
 QY 73 ValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLysLysAsnArgIleIle 92
 Db 208 GCAGCATTAGATGAACCTGTTGTTATTCAGATGATGGCAGCACTAAAAATATAACA 267

RESULT 12
 ACF74415

QY 93 LysProLeuArgAspLysValleGlyLysTyrThrSerValGlyThrArgLysGlnProAsn 112
 Db 268 AAGTCAGTAGAGATAGATTGGGGCATATGAATCGGTGTGATCTAGACCGCAACCGAAT 327
 QY 113 LeuGluGluLeuSerLysLeuLysProAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 132
 Db 328 ATGGAAGTGATAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 387
 QY 133 GlyIleTyrLysAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 152
 Db 388 AAAATCAATCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 447
 QY 153 AspTyrAsnGluAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 172
 Db 448 GATTATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 507
 QY 173 GluGlyLysLysArgLeuGluGluHisAspLysLysLeuGluGlyLysLysLysLys 192
 Db 508 GAAGGCGAGAGCGTCTGGAAAGCATGATAAATATTAATTAATTAATTAATTAATTAAT 567
 QY 193 ThrMetAspLysAsnGlnLysValLeuProAlaValAlaAlaLysSerGlyLeuLeuAla 212
 Db 568 GAACAGAGTACGTTAAATCTGCATTTGCAATTCGATCTCAAGACGAGTATGTTTATT 627
 QY 213 HisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232
 Db 628 AATAATGAGATACATTTATGGACATCTTAATTAATTAATTAATTAATTAATTAATTAAT 687
 QY 233 SerAspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsn 252
 Db 688 ACAAAAGACAAAACACGATGTTGGTGAACGACGAGTGTCTTATATATATTAATTAAT 747
 QY 253 ThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn---LysAla 271
 Db 748 AATGAAGAACTTGCCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 807
 QY 272 SerSerAsnGluProSerLeuLysGluLeuGluLysAspPro---ValTrpLysLysLeu 290
 Db 808 GACAAAATAGAACGAATTC-----ATTGATCTCTGAGTGTGGAAATCAATTA 855
 QY 291 AsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArg 310
 Db 856 AAAGCTGTGAAGATAACAAAGTTATGACGTTGACCGAATAAGTGTGAAATCAAGG 915
 QY 311 GlyLeuLeuSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328
 Db 916 GGGATTATCGCAAGTGAAAGTATGCGCAGAGATTTAGAAAAAATTCGAGAAAAA 969
 RESULT 13
 ID ACC48534 standard; DNA; 2957 BP.
 AC ACC48534;
 XX 11-AUG-2003 (first entry)
 XX 11-AUG-2003 (first entry)
 DE Staphylococcal surface-exposed immunogenic polypeptide DNA.
 XX Surface-exposed immunogenic polypeptide; SEIP; siderophore; receptor;
 KW antibacterial; vaccine; gene; ds.
 XX Staphylococcus aureus.
 OS
 XX
 FH Key Location/Qualifiers
 CDS 983..1960
 FT /*tag= a
 FT /partial
 FT /product= "SEIP"
 FT /note= "No start codon"
 XX
 PN WO2003020875-A2.
 XX
 XX 13-MAR-2003.

XX 17-JUN-2002; 2002WO-US019224.
 XX 17-JUN-2001; 2001US-0298975P.
 XX (DSQU-) D-SQUARED BIOTECHNOLOGIES INC.
 XX Scott DL;
 XX WPI; 2003-300870/29.
 DR P-PSDB; ABR41867.

Novel surface-exposed immunogenic polypeptide of *Staphylococcus aureus* containing receptors for siderophores or iron-binding ligands, useful for producing antibodies effective against *Staphylococcal* infection.

Claim 4; Page 55-56; 62pp; English.

The present sequence is that of DNA encoding novel *Staphylococcus aureus* surface-exposed immunogenic polypeptide (SEIP) D2 SA05. To isolate SEIP genes, an expression library of *S. aureus* genomic DNA was screened with anti-SEIP antisera. DNA was isolated from clones expressing SEIPs, and amino acid sequences were deduced and analysed for conserved and/or functional domains. The SEIP encoded by the present DNA sequence has sequence homology with the siderophore family of periplasmic binding proteins. The invention provides methods for the identification, production and recovery of SEIPs. The SEIPs can be used individually, or in combination, to produce anti-staphylococcal antibodies useful in passive or active immunisation strategies to prevent or contain *staphylococcal* infection. They can also be used to develop diagnostic assays

Sequence 2957 BP; 1160 A; 381 C; 512 G; 904 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5,49e-60 Length: 2957
 Score: 724.50 Matches: 149
 Percent Similarity: 66.67% Conservative: 63
 Best Local Similarity: 46.86% Mismatches: 95
 Query Match: 42.87% Indels: 11
 Gaps: 4

US-10-724-972A-6352 (1-335) x ACC48534 (1-2957)

QY 13 ValIleGlyLeuPheValLeuIleAlaThrAlaLysGlyAsnSerSerSer 32
 Db 1025 GTTGTGTTTATGCTTAATCTTAGTTGTGACAGTAGCGGTGTCGTCACAAAGATACT--- 1081
 QY 33 AsnSerSerLysGluSerSerLysAspGlyValGluIleLysHisGluGlyThrThr 52
 Db 1082 -----GAAGAGAAACTGAAATGACGACAAATAAAGATGAATTAGGAACCGAA 1129
 QY 53 LysValProLysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeu 72
 Db 1130 AAAATTAAGAAAAATCCTAAACGTTGTTGTTATAGATATAGTTTCTGCTGATTATTTA 1189
 QY 73 ValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLysLysAsnArgIle 92
 Db 1190 GCAGCATTAGATATGAACCTGTTGTTATGACAGATGATGACGACCTATAAATAATAACA 1249
 QY 93 LysProLeuArgAspLysLysLysLysTyrThrSerValGlyThrArgLysGlnProAsn 112
 Db 1250 AAGTCAGTAAAGATAGATTGGGGCATATGAATCGGTGTGATCTAGACCGCAACCGAAT 1309
 QY 113 LeuGluGluLysSerLysLeuLysProAspLeuIleAlaAspAsnAsnArgHisLys 132
 Db 1310 ATCGAAGTATAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1369
 QY 133 GlyIleTyrLysAspLeuAsnLysLysLysLysLysLysLysLysLysLysLysLys 152
 Db 1370 AAAATCAATCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1429
 QY 153 AspTyrAsnGluAsnLeuAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 172

||||| 1430 GATTATAATGCAATATTAAACAGTTCCTTAAGCAGTACGCAAGAGAA 1489
173 GluGlyLysLysArgLeuGluGluHisAspLysLysLysGluLys 192
1490 GAAGGCGAGAGCGTCTGGAAGAGCATGATAAATATTAGCGGAGATTAGAAAGAAAT 1549
193 ThrMetAspLysAenGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAla 212
1550 GAACAGAGTACGTTAAATCTGCATTTCGATTCGATTCCTCAAGACGAGTATGTTATT 1609
213 HisProSerAsnSerTyrrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232
1610 AATAATGAAGATACATTTATGGGACAAATCTTAATAAATGGGTATTCAACCTGAAGTC 1669
233 SerAspAspValThrLysGlyLeuSerLysTyrrLysGlyProTyrrLeuGlnMetAsn 252
1670 ACAAAGACAAACTACGATGTTGGTGAACGAGGCGTCTCTATATATATATTTAAAT 1729
253 ThrGluThrLeuSerGlnValAenProGluArgMetPheLeuMetThrAen---LysAla 271
1730 AATGAAGAACTTGCCAAATATCAATCCAAAGATTATGATTTTAGCCACTGACGGAACG 1789
272 SerSerAenGluProSerLeuLysGluLysLysPro---ValTrrLysLysLeu 290
1790 GACAAATAAGAACGAAATTC-----ATTGATCTCGAGTTTGGAAATCATTA 1837
291 AsnAlaValLysAenGlnAtgValAspLeuAspArgLeuTrrPalaArgSerArg 310
1838 AAGCTGTGAAGATAACAAGTTTATGACGTTGACCGAATAAGTTGTTGAAATCAAGG 1897
311 GlyLeuLeuSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328
1898 GGGATTATCGCAAGTGAAGTATGCGAAGATTATAGAAAAAATTCAGAAAAA 1951
RESULT 14
AAT80398/C
ID AAT80398 standard; cDNA; 2247 BP.
XX AC AAT80398;
DT 14-APR-1998 (first entry)
DE Staphylococcus aureus Gene #4 encoding cDNA sequence 1.
XX Staphylococcus aureus WCUH 29; antagonist; antibacterial; immunogen;
KW vaccine; disease; protection; isolation; ss.
XX Staphylococcus aureus.
OS WO9731114-A2.
PN 28-AUG-1997.
XX 25-FEB-1997; 97WO-GB000524.
XX 26-FEB-1996; 96GB-00004045.
XX (SMIK) SMITHKLINE BEECHAM PLC.
XX Burnham MKR, Hodgson JE;
XX WPI, 1997-435166/40.
XX New Staphylococcus aureus polynucleotide and polypeptide(s) - for
PT isolating antagonist of the polypeptide(s) useful as anti-bacterials.
XX Claim 4; Page 30-31; 117pp; English.
XX The present sequence encodes a novel polypeptide, which is optionally
CC expressed in NCIMB 40771. The polypeptide, and polynucleotide encoding
CC it, are derived from Staphylococcus aureus. Cells expressing ligands
CC binding the polypeptide can be used to isolated candidate compounds that

CC bind and inhibit the activity of the polypeptide. Such compounds can be
CC used as anti-bacterial compounds. The polypeptide may also be used as an
CC immunogen to vaccinate an animal for protection against Staphylococcus
CC aureus caused disease
XX Sequence 2247 BP; 704 A; 397 C; 307 G; 802 T; 0 U; 37 Other;
SQ

Alignment Scores: 8.49e-60 Length: 2247
Pred. No.: 721.00 Matches: 169
Score: 73.48% Conservative: 36
Percent Similarity: 60.57% Mismatches: 69
Best Local Similarity: 42.66% Indels: 9
Query Match: 2 Gaps: 1
DB:

US-10-724-972a-6352 (1-335) x AAT80398 (1-2247)

Qy 1 GlyValGluSerValArgGlyLeuLysLysLeuSerValLysGlyLeuLeuPheValLeu 20
Db 831 GGAGTGGGACNGATGAGAGGCTTAAACCTTTAGTATATATTGGGATTAATATTGGCTTAA 772
Qy 21 IleAlaThrAlaAlaCysGlyAenAenSerSerAen-SerSerLysGluSerSerLy 40
Db 771 CTTTATTGTCAGCTTGTGTGTAATACGGATAATCAAGTAAAGAAAGAAATCATCAACTAA 712
Qy 40 sAspGlyVal-GluLysHisGluGluGlyThrThrLysValProLysHisProLysA 60
Db 711 AGATACTATTTCGGTAAAGATGAAATCGTACAGTAAAGTACCTAAAGATGACCAAC 652
Qy 60 rgValValValLeuGluTyrrSerPheValAspAla---LeuValAlaLeuAspValLysp 79
Db 651 GTATCGTTGTATTAGAGTACTCTTTTGCAGGTTGTCATTANCAGCATATACCGCTTANAC 592
Qy 79 ro-ValGlyLysLysAspAenLysLysAenArgLysLysProLysProLysArgAspLys 98
Db 591 CAAGTTGGTATTGCTGATGATGTTGAAGAAAAGAGTATCATAC-CCAGTAAAGAGAAAA 533
Qy 99 IleGlyLysTyrrThrSerValGlyThrArgLysGlnProAenLeuGluGluLysSerLys 118
Db 532 TTGGGGGT-AATACTTCTGTAGGTACACGTA--CAAGCNAACTTTTAAGAGANATTAGTAAA 475
Qy 119 LeuLysProAspLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 138
Db 474 TTAANCCGGGATTAATTTATCNCGTAGTAGTACATAGGTTATTAATTAAGAAATTA 415
Qy 139 AsnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 158
Db 414 GCCCAGATTGACACACA-TTATCATTAAGAGTTTGTGATGGAGACTACACCCACAATATT 356
Qy 159 AspAlaPheLys-ThrLysSerLysAlaLeuGlyLysGluGluGluGlyLysLysArgLys 178
Db 355 AATTGCTTCAACAACAATTTGCTAAAGCTTTAAATAAGANAAAGAGCGNNAAACGCTT 296
Qy 178 uGluGluHisAspLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 198
Db 295 TGCTGAGCATGATAAATAATCAATAGTATTAAGATGAATAATTAATTTGATAGAAATCA 236
Qy 198 nLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTy 218
Db 235 CAAAGTGTCTCCAGCAGTAGTTGCTTAAGCTGGTTTATAGCACATCCACATTTTCA 176
Qy 218 rValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAspValThrly 238
Db 175 TGTGGACAAATTTTAAACGANCATAGGATTTAAATAATGCAATTAAGTGACGATGTACACA 116
Qy 238 sGlyLeuSerLysTyrrLeuLysGlyProTyrrLeuGlnMetAenThrGluThrLeuSerG 258
Db 115 AGGTTTAAAGTAAATATTGAAAGGACCTTACTTACAAATTAGACACTGACATTTAGCTGA 56
Qy 258 nValAenProGluAenMetPheLeuMetThrAenLysAlaSerSerAsn 274
Db 55 TTTAAATCCAGAGCGGTATGATCATCTTATGACAGATCATGCTTAAAAAAGAT 7

RESULT 15
AAT83786/c
ID AAT83786 standard; DNA; 2247 BP.
AC AAT83786;
XX
DT 16-JUL-1998 (first entry)
XX
DE DNA encoding 3 Staphylococcus aureus proteins of unknown function.
XX
KW Staphylococcus aureus protein; ribozyme; antisense sequence; control;
KW Staphylococcus aureus gene; regulatory element; bacterial gene expression;
KW vaccine; Staphylococcus aureus infection; food poisoning; scaled skin syndrome;
KW toxic shock syndrome; ss.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT CDS complement (1212..1445)
FT FT /*tag= a
FT CDS complement (1456..1587)
FT FT /*tag= b
FT CDS complement (1705..2016)
FT FT /*tag= c
XX
XX WO9730070-A1.
XX
XX
XX 21-AUG-1997.
XX
PF 19-FEB-1997; 97WO-US002318.
XX
PR 20-FEB-1996; 96US-0011888P.
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO;
PI Pratt JM, Reichard RW, Rosenberg M, Ward JM;
XX
DR WPI; 1997-424969/39.
XX
DR P-PSDB; AAW27819, AAW27820, AAW27821.
XX
XX Novel polypeptide(s) from Staphylococcus aureus strain WCUH29 - used to
PT isolate antimicrobial compounds, and in vaccines against S. aureus
PT infection.
XX
PS Claim 9; Page 672-673; 989pp; English.
XX
CC The present sequence encodes 3 Staphylococcus aureus proteins of unknown
CC function. The present sequence was isolated from a library of clones of
CC S. aureus WCUH 29 in Escherichia coli. The DNA sequence can be used in
CC the construction of ribozymes and antisense sequences to control the
CC expression of Staphylococcus aureus genes. The DNA sequence is also useful as a
CC source of regulatory elements for the control of bacterial gene
CC expression. The encoded protein may be used to produce vaccines to enable
CC a host to produce specific antibodies with antibacterial action. These
CC vaccines and antibodies would protect a host against invasion by S.
CC aureus, and conditions relating to Staphylococcus aureus infection, e.g.
CC Staphylococcus aureus food poisoning, scaled skin syndrome, and toxic shock
CC syndrome
XX
SQ Sequence 2247 BP; 704 A; 397 C; 307 G; 802 T; 0 U; 37 Other;
XX
Alignment Scores:
Pred. No.: 8.49e-60 Length: 2247
Score: 721.00 Matches: 169
Percent Similarity: 73.48% Conservatives: 36
Best Local Similarity: 60.57% Mismatches: 69
Query Match: 42.66% Indels: 9
DB: 2 Gaps: 1
US-10-724-972A-6352 (1-335) x AAT83786 (1-2247)
Qy 1 GlyValGluSerValArgGlyLeuLysArgLeuSerValIleGlyLeuPheValLeu 20

Db 831 GGAGTGGGACGATGAGAGCTCTAAAAACCTTTTGTATATATGGGATTAAGTGGCTTA 772
Qy 21 IleAlaThrAlaAlaCysGlyAsnAsnSerSerSerSerSerSerSerSerSerSerSerLy 40
Db 771 CTTTGTAGTGGAGCTTGTGTGTAATACGATAATCAAGTAAAGAAAGAAATCATCACTAA 712
Qy 40 sAspGlyVal-GluIleLysHisGluGluGlyThrThrLysValProLysHisProLysA 60
Db 711 AGATCTATTTCGGGTAAAGATGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 652
Qy 60 rgValValValLeuGluGlyThrSerPheValAspAla---LeuValAlaLeuAspValLys 79
Db 651 GTATCGTTGTATTAGAGTACTCATTTCGAAGGTTGCATTATCAGCATATAGCGGTTANAC 592
Qy 79 ro-ValGlyIleAlaAspAspAsnLysLysAsnArgIleLysLysProLeuArgAspLys 98
Db 591 CAAGTTGGTATTGCTGATGATGTAAGAAAAACGATATCATAC-CAGTAAGAGAAAAA 533
Qy 99 IleGlyLysThrThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLys 118
Db 532 TTGGGGGT-AATACCTCTGTAGTACACGTAA-CAAGCNAACCTTAAGAGANATTAGTAA 475
Qy 119 LeuLysProAspLeuIleAlaAspAsnAsnArgHisLysGlyIleLysLysAspLeu 138
Db 474 TTAANCCGGATTAAATTATCTGATAGCAGTAGACATAGGTTAATAAAGAAATTA 415
Qy 139 AsnLysIleAlaProThrIleGluLysSerPheAspGlyAspThrAsnGluAsnIle 158
Db 414 GCCCATTGTCACCA-CATTATCAATAAGAGTTTGTGAGACTACACCCACATATT 356
Qy 159 AspAlaPheLys-ThrIleSerLysAlaLeuGlyLysGluGluGlyLysLysArgLe 178
Db 355 AATTGCTTCAACAACAATTTGCTAAAGCTTTAAATAAGANAAGAAAGCGCNAACGCT 296
Qy 178 uGluGluHisAspLysLysIleGluGluLysLysGluIleThrMetAspLysAsnGlu 198
Db 295 TGCTGAGCATGATAATAATCAATAGTATTAAGATGAATTAATTTGATAGAAATCA 236
Qy 198 nLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTy 218
Db 235 CAAAGTGTCTCCAGCAGTAGTTGCTAAAGCTGTTTATTAGCACATCCACATATTCTA 176
Qy 218 rValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrly 238
Db 175 TGTGGACAATTTTAAACGAGCTAGGATTTAAATAATGCATTAGTACGATGAACACA 116
Qy 238 sGlyLeuSerLysThrLeuLysGlyProTyLeuGlnMetAsnThrGluThrLeuSerCl 258
Db 115 AGGTTTAAAGTAAATTTTAAAGGAGCCTTACTTCAATTAGACACTGAACATTAGCTGA 56
Qy 258 nValAsnProGluArgMetPheIleMetThrAsnLysAlaSerSerAsn 274
Db 55 TTTAAATCCAGAGCGGTATGATCATTTATGACAGATCATGCTTAAAAAAGAT 7
RESULT 16
ADF30765/c
ID ADF30765 standard; DNA; 5718 BP.
XX
AC ADF30765;
XX
DT 12-FEB-2004 (first entry)
XX
DE Bacillus subtilis strain MB1510 integration region DNA SEQ ID NO:20.
XX
KW expression library; Gram-positive bacterium host cell;
KW non-replicating linear integration cassette; gene; Bacillus; ds.
XX
OS Bacillus subtilis.
XX
XX WO2003095658-A1.
XX
PD 20-NOV-2003.

XX PF 07-MAY-2003; 2003WO-DK000301.
XX PR 07-MAY-2002; 2002DK-00000682.
XX PA (NOVO) NOVOZYMES AS.
XX PI Bjornvad ME, Jorgensen PL, Hansen PK;
XX WI MPI; 2004-012127/01.
XX
XX Generating an expression library of polynucleotides by introducing the
XX linear integration cassette into the host cell and selecting or screening
XX for host cells that produce the polypeptides of interest.
XX
XX Example 5; SEQ ID NO 20; 55pp; English.
XX
XX The present invention describes a method for generating an expression
XX library of polynucleotides integrated by homologous recombination into
XX the genome of a competent Gram-positive bacterium host cell. The method
XX comprises: (a) providing a non-replicating linear integration cassette;
XX and (b) introducing the linear integration cassette into the host cell
XX and selecting or screening for host cells that produce the polypeptides
XX of interest. The cassette comprises: (a) a polynucleotide encoding one or
XX more polypeptides of interest; (b) a 5' flanking polynucleotide segment
XX upstream of the polynucleotide of (1) and comprising a first homologous
XX region located in the 3' end of the segment; and (c) a 3' flanking
XX polynucleotide segment downstream of the polynucleotide of (1) and
XX comprising a second homologous region located in the 5' end of the
XX segment. The first and second homologous regions are at least 50, 1000
XX or 1500 bp, each of which has a sequence identity of at least 80, 85, 90
XX or 95-100% with a region of the host cell genome. Also described: (1) a
XX non-replicating linear Gram-positive host cell integration cassette; and
XX (2) a method of producing a polypeptide of interest. The method is useful
XX in generating an expression library of polynucleotides integrated by
XX homologous recombination into the genome of a competent Gram-positive
XX bacterium host cell. The present sequence represents a Bacillus subtilis
XX strain MB1510 integration region, which is used in an example from the
XX present invention.
XX
XX SQ Sequence 5718 BP; 1693 A; 1089 C; 1259 G; 1677 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 2,846-55 Length: 5718
XX Score: 679.50 Matches: 149
XX Percent Similarity: 62.78% Conservative: 50
XX Best Local Similarity: 47.00% Mismatches: 103
XX Query Match: 40.21% Indels: 15
XX DB: 12 Gaps: 4
XX
XX US-10-724-972A-6352 (1-335) x ADF30765 (1-5718)
XX
XX QY 8 LeuLysIleLeuSerValIleGlyLeuLeuPheValLeuLeuAlaThrAlaAlaCysGly 27
XX DB 1021 ATTGCCATCATGAGTGTGTTA---TTGCTCGCTGCTCATGTTATCCGGTGTTCATCA 965
XX
XX QY 28 AsnAsnSerSerSerAsnSerSerLysGluSerSerLysAspGlyValGluIleLysHis 47
XX DB 964 AGCCAGATAACACCGAGCGGCAAAAGCGAGTCTTAAGGATTCACAGATGATCCATGAC 905
XX
XX QY 48 GluGluGlyThrThrLysValProLysHisProLysArgValValValLeuGluTyrSer 67
XX DB 904 GAAGAGGAAAAACGACAGTAAAGCGCACACCTTAAGCGGTGGTGTGCTTGAAGCTTTCA 845
XX
XX QY 68 PheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLys 87
XX DB 844 TTCTTGATCCGCTTCACATCTCGGATTCACCGGTGGGATTCGAGATGACACAAA 785
XX
XX QY 88 LysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThr 107
XX DB 784 AAAGATATGATTAATAAGAGTGTTCGGCGAGCTCCATT---GATTACATCTGTAGGCACA 728
XX
XX QY 108 ArgLysGlnProAsnLeuGluIleSerLysLeuLysProAspLeuIleAlaAsp 127

DB 727 CGCAGCGAACCAATCTTTGAGGTCTATCATGTTCTTGAAGCCTGATTTATCATCTGCTGAC 668
QY 128 AsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeu 147
DB 667 GCTGAGCGCCATPAAAAACATTTTATAACAGCTGAAAAAATCGCCCGACGATTGAATTA 608
QY 148 LysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAla 167
DB 607 AAAAGCCGTGAAGCGACATATGACGAAACGATCGACAGCTTTTACGACCATTTCTTAAAGCA 548
QY 168 LeuGlyLysGluGluGlyLysArgLeuGluGluHisAspLysIleGluGlu 187
DB 547 TTAATAAAGAAGATGAAGGAAAAAGAGTTGCGAGACACAAAAAATCATCAACGAT 488
QY 188 TyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLys 207
DB 487 CTAAAGCCGAACCTTCCGAAAGATGAACCCGCAACATCGTTCTCGCGGTTCGACGAGCG 428
QY 208 SerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGly 227
DB 427 GATTCCTTCCAGCTTCATACATCATCTATGACGAGAAATCTTTAAATGCTTAGGC 368
QY 228 PheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLysGlyPro 247
DB 367 TTATACACACGCTGTGAAGTCAGAT-----AAGGCC 338
QY 248 TyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMet 267
DB 337 TATCAAGAGGTGACCTTGAAGCAATGAGCAAAATCGATCTGATATTTTGTTCATCTCA 278
QY 268 ThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLysAspProValTyr 287
DB 277 GCCAACGAGCAAAACC-----ATTGTAGATGATGGAAACGAAACCCGCTCTGG 227
QY 288 LysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTyrAla 307
DB 226 AAAAATCTCAAGCGGTGAAAAATGGAACAAGTCTATGATCGGACCGGTGACACTGGACA 167
QY 308 ArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGlu 324
DB 166 AGATTTCAGAGGCATCAAGTCTAGTGAACAAGCCCAAGATGTGCTTAAA 116
XX
XX RESULT 17
XX ACA47237
XX ID ACA47237 standard; DNA; 525 BP.
XX AC ACA47237;
XX DT 19-JUN-2003 (first entry)
XX DE Prokaryotic essential gene #28894.
XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;
XX KW drug design; gene.
XX OS Staphylococcus haemolyticus.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (BLIT-) ELITRA PHARM INC.
XX PI Wang L, zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

Db	241	GGGGTCTGCTGATGACAAACCAAAATCGTATATTATAACCATTAAGAAAAAATTTGGA	300
Qy	101	LysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLys	120
Db	301	GATTATAATCTGTGTGCTCGTAAACAACCCAACTTAGAAGAAATCAGTAATAA	360
Qy	121	ProAspLeuIleIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLys	140
Db	361	CCAGATTTAATTATTCAGATAGTAACAGACACAAAGGTATTTACAAAGAATTAAGTAA	420
Qy	141	IleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAla	160
Db	421	ATTGCTCCAACATTGTAATTAAAAAGTTTGTATGAGATTATAATGATAATATAGTCT	480
Qy	161	PhelysThrIleSerLysAlaLeuGlyLysGluGluGlyLys	175
Db	481	TTTAAACAATTGCCAAAGCTTTAAATAAAGATGATGTAGTCTCAG	525
RESULT 18			
ACA42836			
ID	ACA42836 standard; DNA; 897 BP.		
XX	ACA42836;		
XX	19-JUN-2003 (first entry)		
XX	DE		
XX	Prokaryotic essential gene #24493.		
KW	Antisense; ds; prokaryotic essential gene; cell proliferation;		
KW	drug design; gene.		
XX	Pasteurella multocida.		
XX	WO200277183-A2.		
FN	03-OCT-2002.		
PD	21-MAR-2002; 2002WO-US009107.		
PF	21-MAR-2001; 2001US-00815242.		
XX	06-SEP-2001; 2001US-00948993.		
PR	25-OCT-2001; 2001US-0342923P.		
PR	08-FEB-2002; 2002US-00072851.		
PR	06-MAR-2002; 2002US-0267690P.		

(ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

WPI: 2003-029926/02.
P-PSDB; ABU38966.

New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

Claim 14; SEQ ID NO 30706; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation.

Db 70 GTACGGTACAGGATGAGCGGCGAGTTTACCTCAATACCGTTCCCGAGCGGTAGTC 129
Qy 63 ValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIle 82
Db 130 GTGCTGGATTGTCTTCGCGGATGATGGCGGCATCAATATCAGTCCGGTAGGTATT 189
Qy 83 AlaAspAsnLysLysAsnArgHisLysGlyIleTyrLysAspLysIleGlyLysTyr 102
Db 190 GCCGATGATAACGATCCGCGAGCGCATCTCCAGGATGTTTCGCGAGCGTATTAAAGCCCTGG 249
Qy 103 ThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLysLeuLysProAsp 122
Db 250 CAATCAACCGGCGACTCGCGCCCAACAGCCTTGAAGCCATCAGCGCGTGAACCTGAT 309
Qy 123 LeuIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAla 142
Db 310 CTATCATTTGCAGACAGTCAACGCCAGCGGGGATTATCGGCGCACTCAAGGGCATCGCT 369
Qy 143 ProThrIleGluLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLys 162
Db 370 CCGTCTGCTGCTAAATCCCGAACGAAACCTACGAAGAGAACCTGCAATCTGCCGCG 429
Qy 163 ThrIleSerLysAlaLeuGlyLysGluGluGluLysArgLeuGluGluHisAsp 182
Db 430 ATCATCGGCAAGTCATGGGGAAGAGAGTGAAGTGCAGAAACCGCTGGCAGACCGT 489
Qy 183 LysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuPro 202
Db 490 GAACGCAATGAAGCTATGCCAGCGAG-----TTGCCACAAGCGCTCAGTGTGTTT 543
Qy 203 AlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe 222
Db 544 GGCACCTCGCGCAACAGCAATTTAACTGCACCTCCAGCGATACCTATACCGGCGGTC 603
Qy 223 LeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAspValThrLysGlyLeuSerLys 242
Db 604 TTAACAGCATTTAGGGCTAAAGTTCCCGCATCAGTTAAAT----- 642
Qy 243 TyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGlu 262
Db 643 -----CATGGCGCATGTTCTCTCAACCTGGACAGTACTGGCACTCAATCCTGAC 696
Qy 263 ArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGlu 282
Db 697 TGGCTTATCGTGGCCCACTACTCGT-----CAGGAAAGCATCGTGAAACGCTGGCAG 747
Qy 283 LysAspProValTyrLysLysLeuAsnAlaValLysAsnGlnAtgValAspIleLeuAsp 302
Db 748 CAGGATACTCTGTGGCAGATGATGACAGCAGCAGCAGCATCAGATAGCAGCGGTAGAC 807
Qy 303 ArgAspLeuTyrAlaArgSerArgGlyLeuIleSerSerGluLeuMetAlaLysGluLeu 322
Db 808 AGCAACACCTGGGCTCGTATGCGCGGATATTTGGCGCTGAACGCATTTGGCAGCATGCG 867
Qy 323 ValGluLeu 325
Db 868 GTAAAAATC 876
RESULT 21
ID ADR20887/c
ADR20887 standard; DNA; 25860 BP.
AC ADR20887;
XX
XX
XX 07-OCT-2004 (first entry)
XX Photorhabdus luminescens genomic DNA encoding two insecticide proteins.
XX insecticide protein; transgenic plant; insect-resistant plant; wheat;
KW maize; gene; ds.
XX Photorhabdus luminescens subsp. laumondii.

XX Key Location/Qualifiers
PH CDS 20872. .21306
FT /*tag= a
FT /product= "Photorhabdus luminescens insecticide protein
FT #1"
FT 21345. .22598
FT /*tag= b
FT /product= "Photorhabdus luminescens insecticide protein
FT #2"
XX WO2003087377-A1.
PN 23-OCT-2003.
XX 17-APR-2003; 2003WO-FR001239.
XX 17-APR-2002; 2002FR-00004798.
XX (INSP) INST PASTEUR.
PA (CNRS) CNRS CENT NAT RECH SCI.
PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
XX WPI; 2003-833739/77
DR P-PSDB; ADR20889, ADR20891.
XX New nucleic acid from Photorhabdus luminescens, useful for producing
PT insecticidal polypeptides active against Lepidoptera, and for producing
PT insect resistant transgenic plants.
XX Claim 22; SEQ ID NO 1; 44pp; French.
XX The invention comprises the amino acid and coding sequences of proteins
CC from Photorhabdus luminescens which are active against insects. The
CC proteins of the invention are toxic to Lepidoptera, specifically the
CC genera Plutella, Heliothis, Helicoverpa, Spodoptera and Ostrinia. The DNA
CC and protein sequences of the invention are useful in the preparation of
CC transgenic, insect-resistant plants, specifically wheat and maize. The
CC present Photorhabdus luminescens genomic DNA contains the coding sequence
CC for the two insecticide proteins of the invention.
XX Sequence 25860 BP; 7315 A; 5392 C; 6047 G; 7106 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 9.31e-35 Length: 25860
Score: 476.00 Matches: 102
Percent Similarity: 57.24% Conservative: 60
Best Local Similarity: 36.04% Mismatches: 107
Query Match: 28.17% Indels: 14
DB: 11 Gaps: 3
US-10-724-972A-6352 (1-335) x ADR20887 (1-25860)
Qy 43 ValGluIleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValVal 62
Db 8507 GTACGGTACAGGATGAGCGGCGAGTTTACCTCAATACCGTTCCCGAGCGGTAGTC 8448
Qy 63 ValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIle 82
Db 8447 GTGCTGGAAATGCTCATTTGCCGATGATGGCGGCATCAATATCAGTCCGGTAGGTATT 8388
Qy 83 AlaAspAsnLysLysAsnArgHisLysGlyIleTyrLysAspLysIleGlyLysTyr 102
Db 8387 GCCGATGATAACGATCCGCGAGCGCATTTCTACCGATGTTTCGCGAGCGTATTAAAGCCCTGG 8328
Qy 103 ThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLysProAsp 122
Db 8327 CAATCAACCGGCGACTCGCGCCCAACAGCCTTGAAGCCATCAGCGCGCTGAACCTGAT 8268
Qy 123 LeuIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAla 142
Db 8267 CTCATCATTTGCAGACAGTCAACGCCAGCGGGGATTTATCGGCGACTCAAGGGCATCGCT 8208

[illegible]

ACF67367_39/c
Continuation (40 of 57) of ACF67367 from base 3900001 (Photorhabdus luminescens nucleotide sequence) split into 57 fragments
WP Sequence split into 57 fragments
LOCUS ACF67367 Accession ACF67367

Qy	63	VallLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspVallysProValGlyIle	82
Db	95057	GTGCTGGAAATTGTCATTTGCCGATGATGGCGGCATCAATATCATGCTCCGTAAGTATT	94998
Qy	83	AlaAspAspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyr	102
Db	94997	GCCGATGATAAGTATCCGACGGCATTCTACCGATGTTCCGAGCGATTAAAGCCCTGG	94938
Qy	103	ThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLysProAsp	122
Db	94937	CAATCAACCGGCACCTCGCGCCCAACCAACGCTTTGAAGCCATTCAGCGCGTGAACCTGAT	94878
Qy	123	LeuIleIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAla	142
Db	94877	CTCATCATTTGCAGACAGTCAACGCCACCGCGGGATTTATCGGCACCTCAAGGSCATCGCT	94818
Qy	143	ProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLys	162
Db	94817	CCTGTCGTGCTGTAAATTCGCAACGAAACCTACGAAGAGAACTCTGCAATCTGCCGCG	94758
Qy	163	ThrIleSerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAsp	182
Db	94757	ATCATCGGCAAAAGTCATGGGGAAGAAGAGTAGATGTCAGAAACGCGTGCAGAGCACCGT	94698
Qy	183	LysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuPro	202
Db	94697	GAACGCATGAAGGCTATCCAGCCAG-----TTGCCACAAGGGCTCAGTGTGGTTTTT	94644
Qy	203	AlaValAlaIatLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGluGlnPhe	222

US-10-724-972A-6352 (1-335) x ACF67367 39 (1-110000)

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR P-PSDB; ABU28130.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 19870; 1766pp; English.
XX

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 900 BP; 198 A; 266 C; 256 G; 180 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4, 95e-35 Length: 900
Score: 460.00 Matches: 101
Percent Similarity: 57.09% Conservative: 60
Best Local Similarity: 35.82% Mismatches: 105
Query Match: 27.22% Indels: 16
DB: 8 Gaps: 4

US-10-724-972A-6352 (1-335) x ACA32000 (1-900)

Qy	45	IleLysHisGluGluGluThrLysValProLysHisProLysArgValValLeu	64
Db	70	GTTCAGGACGACACGGACGTTTACATCTGATAAAGCCACACGAGTGTGTGCTG	129
Qy	65	GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp	84
Db	130	GAATCTCTGTCGCGATGCGTGGCCGCGTGGACGTCAGCCGATCGTATTGCCGAC	189
Qy	85	AspHisLysLysAsnArgIleLysProLysArgAspLysIleGlyLysTyrThrSer	104
Db	190	GATACCATGATCAAAACGATCTCTGCGAAGTGGTGGCCACCTGAAACCGTGGCAGTCC	249
Qy	105	ValGlyThrArgLysGlnProAsnLeuGluLysSerLysLeuLysProAspLeuIle	124
Db	250	GTCGAGACGCGCGCAGCCGATGGAAGCCATTCCTGCTCCCTGAAACACGACCTGATC	309
Qy	125	IleAlaAspAsnAsnArgHisLysGlyLysLysAspLeuLysLysIleAlaProThr	144

Db	310	ATTGCCGACAGCAGTCGCCATGCGGGATTTACATCGCTTGCAGCAAAATCGCGCCGGTA	369
Qy	145	IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle	164
Db	370	CTGCTGTTAAAGTCCCGCAACGAACTACGCTGAAATTTGCACTCAGCGGCTATCATC	429
Qy	165	SerLysAlaLeuGlyLysGluGluGlyLysArgLeuGluGluHisAspLysLys	184
Db	430	GCGAAGTGTGATGATAAAACGAGAGATGACAGCGCTCTGGAACAACATAAGAGAAG	489
Qy	185	IleGluGluTyrLysLysGluLysGluLysMetAspLysAsnGlnLysValLeuProAlaVal	204
Db	490	ATGGCGCAGTGGTCCAGCCAG-----CTTCCCACAGGACACGCGTGGCTTTGGCACA	543
Qy	205	AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer	224
Db	544	TCACGCGAAGCAGCTATTCAACCTGCATACCCAGGACACCTGGACCGCGCTGGCT	603
Qy	225	GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu	244
Db	604	TCTCTGGGCTG-----AACGTTCCCGCTGGCATG	633
Qy	245	LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg	263
Db	634	GCGGCGCGTCCATGCGCTCCATCGGCTGGAGCAACTGCTGCGGTCAATCTCGCTGG	693
Qy	264	MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLys	283
Db	694	CTGTGGTGGCCCACTATCGC-----GAAGAGAGCATTTGTTAAACGCTGGCAACAA	744
Qy	284	AspProValTyrLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg	303
Db	745	GATCGCTCTGGCAGATATTACCGCGCGCAGAGAAGCAGCAGGTTCCTTCAGTCGACAGT	804
Qy	304	AspLeuTyrAlaArgSerArgGlyLeuLysSerSerGluMetAlaLysGluLeuVal	323
Db	805	AACGCTGGCGCGGATGCGGTATTTTTCGCGAGAGCGGTATTGCGCTCACACGGTA	864
Qy	324	GluLeu 325	
Db	865	AAAATC 870	
RESULT 25			
AAS52735			
ID	AAS52735	standard; DNA; 909 BP.	
XX			
AC	AAS52735;		
XX			
DT	13-FEB-2002	(first entry)	
XX			
DE	E. coli	DNA for cellular proliferation protein #457.	
XX			
KW	Antisense; ds;	prokaryotic cellular proliferation gene; antibiotic;	
KW	antibacterial;	drug design.	
OS	Escherichia coli.		
XX			
FN	WO200170955-A2.		
XX			
PD	27-SEP-2001.		
XX			
PF	21-MAR-2001;	2001WO-US009180.	
XX			
PR	21-MAR-2000;	2000US-0191078P.	
PR	23-MAY-2000;	2000US-0206848P.	
PR	26-MAY-2000;	2000US-0207727P.	
PR	23-OCT-2000;	2000US-0242578P.	
PR	27-NOV-2000;	2000US-0253625P.	
PR	22-DEC-2000;	2000US-0257931P.	
PR	16-FEB-2001;	2001US-0269308P.	
XX			
PA	(ELIT-) ELITRA	PHARM INC.	
XX			

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
DR WPI; 2001-611495/70.
DR P-PSDB; AAU34876.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Claim 27; SEQ ID NO 6372; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence encodes an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 909 BP; 200 A; 262 C; 261 G; 186 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,01e-35 Length: 909
Score: 460.00 Matches: 100
Percent Similarity: 57.45% Conservative: 62
Best Local Similarity: 35.46% Mismatches: 104
Query Match: 27.22% Indels: 16
DB: 4 Gaps: 4

US-10-724-972A-6352 (1-335) x AAS52735 (1-909)

QY 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64
DB 76 GTTCAGGACGAAACACGGACGTTTACATCTCGAATAAAGCCACACGGATGTTGGTGTG 135
QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84
DB 136 GAACCTCTGTCGCGATGCGTGGCGCGCGGACGTCATCCGATCGGTATTGCCGAC 195
QY 85 AspAsnLysLysAsnArgIleLysProLeuArgAspLysIleGlyLysThrSer 104
DB 196 GATAACGATGCAAAACGATCTGCGCGAAGTGGCGCACCTGAAACCGGTGGCGAGTCC 255
QY 105 ValGlyThrArgLysGlnProAsnLeuGluGlyLeuSerLysLysLysProAspLeu 124
DB 256 GTCCGAACCGCGCGCGAGCGGACCTGGAAGCCATTCGCGCTGAAACACACCTGATC 315
QY 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspAsnLysIleAlaProThr 144
DB 316 ATTGCCGACGAGCATGCGCCATGCGGGGTTTACATCGCTTGCAGCAATCGCGCGGTA 375
QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164
DB 376 CTGCTGCTTAAGTCCCGCAACGAACTAGCTGTAATAATTTGCAATTCGCGGTATCATC 435
QY 165 SerLysAlaLeuGlyLysGluGluGlyLysLysLysLysLysLysLysLysLys 184
DB 436 GCGCAATGCTGGGTAAAGACGAGAGATCGACGCGACGTCGTGGAACACATAAGAGAGG 495
QY 185 IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204
DB 496 ATGGCGGACGTGGCCAGCCAG-----CTTCCCAAGGGACACGCGTGGCGCTTTGGCACA 549

QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224
DB 550 TCACGCGAAGACAGCAATTCAACTGATCTAGGAGACCTGGACCGGACGCTGCTGGCC 609
QY 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu 244
DB 610 TCTCTGGGCTG-----AACGTTCCCGCTGCGATG 639
QY 245 LysGlyProTyrIleu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263
DB 640 GCGGGTGGTCCATGCGCTCATCGCTGAGCAACTGCTGCGGTCAATCTGCTGG 699
QY 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLys 283
DB 700 CTCTGCTGTTGCCCACTATCGC-----GAAGAGAGCATTTGTTAAACCTGCGCAACA 750
QY 284 AspProValTrpLysLysLeuAlaValLysAsnGlnArgValAspIleLeuAspArg 303
DB 751 GATCCGCTCTGGCAGATGTTAAACCGCGCGCAGACGAGGTGCTTGGTTCGACAGT 810
QY 304 AspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuVal 323
DB 811 AACACCTGGCGCGGATGCGCGGTATTTGCTGCGAGCGGTATTCGCGTGCACCGTA 870
QY 324 GluLeu 325
DB 871 AAAATC 876
RESULT 26
ACA32751
ID ACA32751 standard; DNA; 909 BP.
XX
AC ACA32751;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #14408.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Escherichia coli.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
XX
DR P-PSDB; ABU28881.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 20621; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:

(1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 909 BP; 200 A; 262 C; 261 G; 186 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5,01e-35	Length:	909
Score:	460.00	Matches:	100
Percent Similarity:	57.45%	Conservative:	62
Best Local Similarity:	35.46%	Mismatches:	104
Query Match:	27.22%	Indels:	16
DB:	8	Gaps:	4

US-10-724-972A-6352 (1-335) x ACA32751 (1-909)

QY	45	IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValLeu	64
DB	76	GTTCAGGACGACACGGCAGCTTACACTCGAATAAAACGCCACGAGTGTGGTGTG	135
QY	65	GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyLeuAla	84
DB	136	GNACTCTGTTCCCGATGCGCTGCGCGCGGTGGACGTCATCCCGATGCGTATTG	195
QY	85	AspAsnLysLysAsnArgIleLysProLysArgAspLysIleGlyLysTyrThrSer	104
DB	196	GATACGATGCAAAACGATCTGCGCGAGTGGCTGGCCACTGAAACCGTGGCAGTCC	255
QY	105	ValGlyThrArgLysGlnProAsnLeuGluGluLysSerLysLeuLysProAspLeu	124
DB	256	GTGGAACCGCGCGCAGCGCCGCTGGAAGCCATTGCGCTCTGAAACACGACCTGATC	315
QY	125	IleAlaAspAsnAsnArgHisLysGlyLeuLysAspLeuAsnLysIleAlaProThr	144
DB	316	ATTGCGACACAGTGTGCGGCGGTGTACATCGCTTGCAGCAATCGCGCGGTA	375
QY	145	IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThr	164
DB	376	CTGCTGCTTAAGTCCCGACAAACCTACCTGAAATAATTTGCAATCTGGCGGTATCATC	435
QY	165	SerLysAlaLeuGlyLysGluGluGlyLysArgLeuGluGluHisAspLysLys	184
DB	436	GGCGAAATGTGTGGTGAAGAGAGATGACGACGCTGTGGAACAACATAAGAGAGG	495
QY	185	IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAla	204
DB	496	ATGCGCGAGTGGCCAGCCAG-----CTTCCCAAGGGACACGCGTGGCTTTGGCACA	549

QY	205	AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer	224
DB	550	TCACGGGACAGCAATTCACCTGATCTCAGGAGACCTGGACCGCGCTGCTGGCC	609
QY	225	GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu	244
DB	610	TCTCTGGGGCTG-----AACGTTCCCGTGGCGATG	639
QY	245	LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg	263
DB	640	GCGGGTGGCTGCTGCGCTGCGCTGCGCACTGCTGGCGGTCAATCTGCTGG	699
QY	264	MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLys	283
DB	700	CTGCTGGTTGGCCACTATCGC-----GAAGAGAGCATTTGTTAAACGCTGGCAACA	750
QY	284	AspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg	303
DB	751	GATCGCTCTGGCAGATGTTACCGCGCGCAGACGACGAGTTCCTCGGTCGACAGT	810
QY	304	AspLeuTrpAlaArgSerArgLysLeuLysSerSerGluGluMetAlaLysGluLeuVal	323
DB	811	AACACCTGGCGCGGATGCGCGGTATTTTGTCTGCAGAGCGTATTGCGCGTACACG	870
QY	324	GluLeu 325	
DB	871	AAAAATC 876	

RESULT 27
AAS46238/c

ID AAS46238 standard; DNA; 10244 BP.

AC AAS46238;

XX 18-DEC-2001 (first entry)

DE DNA encoding novel mar regulated protein (NIMR) #7.

KW mar regulated polypeptide; NIMR; microbial infection; antibacterial; ds.
OS Escherichia coli.

PN WO200170776-A2.
XX 27-SEP-2001.

PF 08-MAR-2001; 2001WO-US007478.
XX 10-MAR-2000; 2000US-0188362P.

PR (TUFT) TUFTS COLLEGE.
XX Levy SB, Barbosa TM, Alekshun MN;
PI WPI; 2001-602769/68.
DR P-PSDB; AAU29336.

XX Identifying compounds that modulate a newly identified mar regulated polypeptide activity, useful as antimicrobial compounds, involves contacting the polypeptide with a test compound.

PS Disclosure; Page 178-184; 526pp; English.

XX The invention relates to a method of identifying compounds that modulate a newly identified mar regulated (NIMR) polypeptide activity. The method comprises contacting an NIMR polypeptide with a test compound under interaction conditions, determining the ability of the compound to modulate the activity or expression of the polypeptide, and selecting the modulators. NIMR nucleic acids and polypeptides are used in the treatment of microbial infections, and in screening for modulators of NIMR expression and activity. These modulators can be used to reduce the infectivity of a microbe on a surface, and the virulence of a microbe in a subject suffering from an infection. AAS46232-AAS46278 represent

RESULT 28
ACA45366
TD ACA45366 standard; DNA; 891 BP.

XX	ACA45366;
AC	
XX	19-JUN-2003 (first entry)
DT	
XX	Prokaryotic essential gene #27023.
DE	
XX	Antisense; ds; prokaryotic essential gene; cell proliferation;
KW	drug design; gene.
KW	

KW drug design; gene.
 XX
 OS *Pseudomonas syringae*.

OS	<i>Pseudomonas syringae</i> .
XX	
PN	WC200277183-A2.

XX
PD
XX
03-OCT-2002.

PF 21-MAR-2002; 2002WO-US0099107.
XX
PR 21-MAR-2001; 2001US-00815242.

06-SEP-2001; 2001US-00946553.
PR
PR
PR 25-OCT-2001; 2001US-034923P.
PR
PR 08-FEB-2002; 2002US-00072851.
PR

PR 06-MAR-2002; 200203-03020551.
XX
PA (ELIT-) ELITRA PHARM INC.

22 Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Pi Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

WPI; 2003-029926/02.
DR
P-PSDB; ABU41496.
DR
324

PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX
PS
XX

Claim 14; SEQ ID NO 33236; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising a sequence
CC of the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC nucleic acid whose expression is inhibited by the

CC polypeptide or its fragment was specifically binding
CC antipeptide; (4) an antibody capable of specifically binding
CC antisense nucleic acid; (5) producing the polypeptide; (6) inhibiting cellular
CC the polypeptide; (7) the gene in an operon required for
CC the synthesis of a gene in an operon required for

[illegible]

identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies as a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC of overexpression or underexpression

to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for

CC identifying proteins or screening for homologous nucleic acids
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids

CC required for proliferation in cells other than *S. aureus*. The present sequence is one of the targets
CC K. pneumoniae or *P. aeruginosa*. Note: The sequence data for this patent did
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC prokaryotic essential genes. Note: The sequence data for this patent did

CC not form part of the printed specification,
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 891 BP; 168 A; 276 C; 284 G; 163 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.86e-34	Length:	891
Score:	454.00	Matches:	110
Percent Similarity:	52.80%	Conservative:	60
Best Local Similarity:	34.16%	Mismatches:	104
Query Match:	26.86%	Indels:	48
DB:	8	Gaps:	7

US-10-724-972A-6352 (1-335) x ACA45366 (1-891)

```

QY 17 LeuPheValLeuIleAlaThrAlaAaCysGlyAsnAsnSerSerSerAsnSerSerLys 36
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 13 CTGCTCAGCGTCTGCGAGCGCGGT----- 39

QY 37 GluSerSerLysAspGlyValGluIleLysHisGluGluGlyThrThrLysVal----- 54
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 40 -----CAGCGTCGCGCCATCGACATCGATCAGCGGCAGCACAGGTGCATCTG 87

QY 55 ProLysHisProIysArgValValLeuGluTyrSerPheValAspAlaLeuValAla 74
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 88 CCAGACACCCCAAGCGCGTGTGGTACTGGAATTTTCGTTTCTCGACGGGCTTGCTCG 147

QY 75 LeuAspValLysProValGlyIleAlaAspAspAsnLysLysAsnArgIleLysPro 94
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 148 GTTGGCGTGACACCGGTTCGCGCCGCGAGATGATGGCGACGCCAGTCGTATCTGCCCAAG 207

QY 95 LeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsnLeuGlu 114
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 208 GTGCGCAAGCGCGTGGGTGAATGGCAGTCGTTGGGGCTCGCTCGCAACCCCAATATCGAA 267

QY 115 GluIleSerLysLeuLysProAspLeuIleIleAlaAspAsnAsnArgHisLysGlyIle 134
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 268 GTGATCGCAGCGCTCAAGCCGAGCTGATCATCGCTGACCTCGGTCTCATCAGGCTCTT 327

QY 135 TyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGlyAspTyr 154
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 328 TATTAACGACCTTGGCCAGCTGTGCTCCGACGCTGATGTTGCCGTCCAGCGGCGAGACTAT 387

QY 155 AsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGly 174
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 388 CAGGCGACCTGAATCTGCGGGCTGATCGGCATCGCGCTGGCGCTGGCAAGGCCGCGAGATG 447

QY 175 LysLysArgLeuGluGluHisAspLysLysIleGluGluTyrLysLysGluIleThrMet 194
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 448 CAGGCGCGAATCGCAGAAAAACCGTCAGCACCTGAAAAACCGTTGCCGACAGATTCTCTGCC 507

QY 195 AspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisPro 214
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 508 GACAGCAAC-----GTGCTCTTCGGCGTTGCCGTGAAGACAGTTTCTCGGTCCACGGC 561

QY 215 SerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAsp 234
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 562 CCGCATTCTTACGCGAGGCGGTGCTGCAAGCCATCGGT----- 600

QY 235 AspValThrLysGlyLeuSerLysTyrLysLysGlyProTyrLeuGlnMetAsnThr--- 253
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 601 -----TTGCAAGTACCGGAAGTGGCGCAACAGCCGCGCG 633

QY 254 -----GluThrLeuSerGlnValAsnGlnArgValAspIleLeuAspArgAspLeu 266
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 634 CCCACCAGTTCTGCAGCGTGGAGCAACTGCTCGCCCTGGACCCCACTGTTGCTGCTGTC 693

QY 267 MetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu---GluLysAspPro 285
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 694 -----GGCCATTATCGTCGCCCGAGCATTTGTCGACACTCGAGCAAGCAGCGCG 741

QY 286 ValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeu 305
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 742 CTGTGCGCAGTGTCTGGCGCTGTACGTAAATAAACAGGTGCGCGAAGTCTGACGGTATGAC 801

```


Alignment Scores:

Pred. No.:	1.19e-29	Length:	999
Score:	405.00	Matches:	113
Percent Similarity:	51.74%	Conservative:	65
Best Local Similarity:	32.85%	Mismatches:	132
Query Match:	23.96%	Indels:	34
DB:	2	Gaps:	12

US-10-724-972A-6352 (1-335) x AAZ22850 (1-999)

QY	8	LeuLysIleLeuSerValIleGlyLeuPheValLeuIleAlaThrAlaAlaCysGly	27
DB	19	ATTAATAAGCTT---GTGGTTACGCTTGCTTTCTACTTGT---TTAGCAGCATGTAGT	72
QY	28	AsnAsnSerSerSerAsnSerSerLysGluSerSerLysAspGlyValGluIleLysHis	47
Db	73	GGAATTCAAATAAACAATCACTGATACAAAGATAAGGAACAACACTTCAATTAACAT	132
QY	48	GluGluGlyThrThrLysValprolyshisProLysArgValValValLeuGluTySer	67
Db	133	GCAATGGGTACACTGAATTAAGGGNAACCAAAGCGTGTGTGTACGTATATCAAGGT	192
QY	68	PheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsnLys	87
Db	193	GCCACTGACGTCGCTGATCTTTAGGTGTTAACTCTAGTCTGTAGAATCATGCACA	252
QY	88	LysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyThrSerValGlyThr	107
Db	253	CAAAAACCGAAATTCGAATACATAAAAAATGATTTAAACATACTAAGATTGTAGTCAA	312
QY	108	ArgLysGlnProAsnLeuGluIleSerLysLeuLysProAspLeuIleAlaAsp	127
Db	313	GAACCTGCACCTAACTTAGAGGAATCTTAATTTAAACCGGACTTAATTGTGCGTCA	372
QY	128	AsnAsnArgHisLysGlyIleTyLysAspLeuAsnLysIleAlaProThrIleGluLeu	147
Db	373	AAAGTTGAATGAAAGATTTCAGATCAATATCTAAAATCGCACCAACAGTT-----	426
QY	148	LysSerPheAppGlyAspTyArgAsnGluAsnIleAspAlaPheLysThrIleSerLysAla	167
Db	427	--TCTACTGATCAGCTTTCAAAATTCAAA--GATACAACCTAAGTTAATGGGAAAGCT	480
QY	168	LeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGlu	187
Db	481	TTAGCGAAAGAAAGAAAGCTGAAGATTACTTTAAAAGTACGATGATTAAGTACTGCA	540
QY	188	TyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLys	207
Db	541	TTCCAAAAA-----GATGCAAAAGCAAGATATAAAGATGCATGGCCATTGAA	588
QY	208	SerGlyLeuLeuAlaHisProSerAsn-----SerTyrValGly	220
Db	589	GCITCAGTTGTAACTTCGCTCGTGCATCATACAAGAAATTATGCTGGTATGCTGCT	648
QY	221	GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu	240
Db	649	GAATCTTAATGATTAGATTCAACGT-----AATAAGACTTA	690
QY	241	SerLysTyrLeuLysGly-----ProTyrLeuGlnMetAsnThr---GluThrLeuSer	257
Db	691	CAAAAAACAAGTTGATAATGTAAGATATATCCAACCTTACATCTAAGAAGAAGCTCCA	750
QY	258	GlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSer	277
Db	751	TTAATGAACGCTGATCATATTTTGTAGTAAATCAGATCCAAATGCCAAGAGTCTGCA	810
QY	278	Leu-----LysGluLeuGluLysAspProValTrpLysLysLeuAsnAla	292
Db	811	TTAGTTAAAAAGACTGAAAGCGAATGGACTTCAAGTAAAGAGTGGAAAAATTTAGACGA	870
QY	293	ValLysAsnGlnArgVal---AspIleLeuAspArgAspLeuTrpAlaArgSerArgGly	311
Db	871	GTTAAAAAACCAAGATATCTGATGATTAGATGAATCACTTGGAACTTAGCTGGCGGA	930

Qy	312	LeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAspSerLys	331
Db	931	TATAAATCTTCATTAATAAACTTTATTGACGATTATATGAA-----AAGTTAAATATTGAA	984
Qy	332	LysAspAsnLys	335
Db	985	AAACAATCAAAA	996

RESULT 31

AAZ19889
ID AAZ19889 standard; DNA, 999 BP.
XX
XX
AC AAZ19889;
XX
XX
DT 06-DEC-1999 (first entry)
XX
XX
DE Staphylococcus aureus iron regulation gene cbrA.
XX
XX CbrA gene; infection; therapy; diagnosis; vaccine; antibiotic;
KW iron regulation,ss.
KW

OS Staphylococcus aureus.

XX	Key	Location/Qualifiers
FH	CDS	7 . .999
FT		/*tag= a
FT		

PN WO9947639-A2.

XX
PD 23-SEP-1999.

XX
PF 19-MAR-1999; 99WO-US005976.

XX
PR 20-MAR-1998; 98US-0078682P.

PR 01-APR-1998; 98US-0080296P.
PR 07-MAY-1998; 98US-0084674P.

XX
PA (HUMA-) HUMAN GENOME SCI INC.

PA (LUDW-) LUDWIG INST CA

XX
PI Simpson AJG, Choi GH;

XX
DR WPI; 1999-562101/47.

DR P-PSDB; AAY31824.
XX

PT New isolated *Staphylococcus* nucleic acid molecules, used to develop
PT products for the diagnosis, prevention and treatment of *Staphylococcal*
PT infections.

PS
PS
Claim 1; Page 24; 102pp: English.

This is the nucleotide sequence of *Staphylococcus aureus* strain ISP3 (ATCC 202108) genomic DNA including the novel cbrA gene that codes for a 330-amino acid protein (see AAY31824) of predicted mol.wt. 36.8 kDa. The sequence was obtained from overlapping clones BTAC444 and BTAGJ54, which span a single operon containing the cbrA, cbrB and cbrC genes (see AAY19889-91). cbrA shows sequence homology to known genes involved in iron regulation. The invention provides 11 novel genes (see AAY19882-92) of *S. aureus* and the polypeptides they encode (see AAY31817-27). Also provided are vectors, host cells, antibodies and hybridomas. The invention further relates to screening methods for identifying agonists and antagonists of *S. aureus* polypeptide activity, and to diagnostic methods for detecting *Staphylococcus* nucleic acids, polypeptides and antibodies in a biological sample. Antagonists of cbrA may be useful as antibiotics to treat infections of *S. aureus* and other *Staphylococcus* spp. Also provided are novel vaccines for the prevention or attenuation of infection by *Staphylococcus*. The isolated nucleic acid molecule is also useful for generating probes and primers, and in the recombinant production of cbrA protein.

Sequence 999 BP; 402 A; 142 C; 180 G; 275 T; 0 U; 0 Other;

Alignment Scores:	1.19e-29	Length:	999
Pred. No.:	405.00	Matches:	113
Score:	51.74%	Conservative:	65
Percent Similarity:	32.85%	Mismatches:	132
Best Local Similarity:	23.96%	Indels:	34
Query Match:	2	Gaps:	12
DB:			
US-10-724-972A-6352 (1-335) x AA219889 (1-999)			
Qy	8	LeuLysIleLeuSerValIleGlyLeuLeuPheValIleLeuAlaThrAlaAlaCysGly	27
Db	19	ATTAAATGTT---GTTGTACGCTGCTTCTACTGTT---TTAGCAGATGTAGT	72
Qy	28	AsnAsnSerSerSerAsnSerSerLysGluSerSerLysAspGlyValGluIleLysHis	47
Db	73	GGGAATTCAAATAAACAATCATCTGATAACAAGATAAGGAACAACCTCAATTAACAT	132
Qy	48	GluGluGlyThrThrLysValProLysHisIleProLysArgValValValLeuGluTyrSer	67
Db	133	GCAATGGGTACAACTGAAATTAAGGGAAACCAAGCGTGTGTTCAGCTATATCAAGGT	192
Qy	68	PheValLeuAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsnLys	87
Db	193	GCCACTGACGTCGCTGATCTTTAGGTGTTAAACCTGTAGTGCTGTAGTAATCATGACA	252
Qy	88	LysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThr	107
Db	253	CAAAAACCGAAATTCGAATACATAAAATAATGATTTAAAGATATAAGATTGTAGGTCAA	312
Qy	108	ArgLysGlnProAsnLeuGluGluIleSerLysLysLysProAspLeuIleAlaAsp	127
Db	313	GAACCTGCACCTAACTTAGAGGAATCTCTAAATTAACCGGACTTAATGTGCGGTCA	372
Qy	128	AsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeu	147
Db	373	AAAGTTAGAAATGAAAAAGTTTACATCAATATCTAAATCGCACCAACAGTT----	426
Qy	148	LysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAla	167
Db	427	---TCTACTGATACAGTTTTCAAATTCAAA---GATACAACTTAAGTTAATGGGAAAGCT	480
Qy	168	LeuGlyLysGluGluGluGlyLysLysArgLeuGluHisAspLysValIleGluGlu	187
Db	481	TTAGGCAAGAAAAAGAGCTGAAGATTTACTTTAAAGAGTAGATGAATAAGTAGCTCCA	540
Qy	188	TyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLys	207
Db	541	TTCCAAAAA-----GATCAAAAGCAAGATATAAGATGCATGGCCATTGAAA	588
Qy	208	SerGlyLeuLeuAlaHisProSerAsn-----SerTyrValGly	220
Db	589	GCTTCAGTTGTTAACTTCGCTGCTGATCATCAAGAATTTATGCTGTGGATATGCTGT	648
Qy	221	GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu	240
Db	649	GAATCTTAAATGATTAGGATTCAAACGT-----AATAAGACATT	690
Qy	241	SerLysTyrLeuLysGly-----ProTyrLeuGlnMetAsnThr---GluThrLeuSer	257
Db	691	CAAAAAACAAGTTGATTAATGGTAAGATATTAATCCAACTTACATCTAAAGAAAGCATTC	750
Qy	258	GlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSer	277
Db	751	TTAATGAACGCTGATCATATTTTGTAGTAAATCATAGATCCAAATCGAAAGATGTGCA	810
Qy	278	Leu-----LysGluLeuGluLysAspProValTrpLysLysLeuAla	292
Db	811	TTAGTTAAAAAGACTGAAGCGAATGACCTTCAAGTAAAGAGTGGAAAAATTTAGACGCA	870
Qy	293	ValLysAsnGlnArgVal---AspIleLeuAspArgAspLeuTrpAlaArgSerArgLys	311
Db	871	GTTAAAAACAACCAAGATCTGATGATTTAGATGAATCATCTGGAATTAGCTGCGCGA	930

[illegible]

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target

CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 972 BP; 405 A; 147 C; 183 G; 237 T; 0 U; 0 Other;

Alignment Scores: 2.6e-28 Length: 972
Pred. No.: 391.00 Matches: 116
Score: 51.93% Conservative: 59
Percent Similarity: 34.42% Mismatches: 122
Best Local Similarity: 23.14% Indels: 40
Query Match: 8 Gaps: 13
DB:

US-10-724-972A-6352 (1-335) x ACA21545 (1-972)

QY 9 LysLeuSerValIleGlyLeuLeuPheValLeuIleAlaThrAlaCysGlyAsn 28
Db |||:|||||:
7 AAATCTCTAGTATT-----TTCATAGTAGTTTCTATTCTCGTGTGGATGC 54
QY 29 AsnSerSerAsnSerSerLysGluSerSerLysAsp-----GlyValGlu 44
Db |||:|||||:
55 GGACAGCAAAAGAGGAGAGAAAGAAAGAAAGAAAGCGGCAATATAAATCAAGCTATTACA 114
QY 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64
Db |||:|||||:
115 ATTAACACCGCTGAAGGGGAAACCAAGTTAGATAAACCAGCGAAAGTAGTGTACTT 174
QY 65 GluTyrSerPheValAlaLeuValAlaLeuValLysProValGlyIleAlaAsp 84
Db |||:|||||:
175 GAATGGGTATTATCAGAAAGACTTATTAGCACTTGTGTTTCAGCCAGTAGGGATGGCAGAC 234
QY 85 AspAsnLysLysAsnArgIleIle-----LysProLeuArgAspLysIleGly 100
Db |||:|||||:
235 ATTAAGAATTATATAATGGTAAATACAAAACAAACCGAGTAAAGATGTTGTA--- 291
QY 101 LysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluLeuSerLysLeuLys 120
Db |||:|||||:
292 -----GATGTCGGGACACGTCACCAACCAACCACTTAGAAGAAATTAGCGCTTAAAA 342
QY 121 ProAspLeuIleIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLys 140
Db |||:|||||:
343 CCAGATTTAATATATACAGCTTCATTCGTTGAAGCAATATAAATGAATAGAACAA 402
QY 141 IleAlaProThrIleGluLeu-----LysSerPheAspGlyAspTyrAsnGluAsn 157
Db |||:|||||:
403 ATTGACCACACAGTTATGTTGATCCATCAACACAGCAATAACGATCATCTTTGCTGAAATG 462
QY 158 IleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGluGlyLysLysArg 177
Db |||:|||||:
463 ACAGAAACATTTAAACAAATTCGAAGACGTTGGAAGAAAGAAAGAAAGGTAAGTA 522
QY 178 LeuGluGluHisAspLysLysIleGluGluTyrLysLysGluIle-----Thr 193
Db |||:|||||:
523 TTAGCTGATATGATGAATAGCATTCGCTGATGCAAGCAAAATTTGCAAAAGCAGACTTA 582
QY 194 MetAspLysAsnGlnLysValLeuProAlaValAlaLysSer-----Gly 209
Db |||:|||||:
583 AAAGATAAAACATCGCAATGGCAGACGATTTACTGCTTAAATATGTCACATTCCT 642
QY 210 LeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLys 229
Db |||:|||||:
643 ATCTTAAC-----GACAACTTTAGCTTTACAAAGTTACAAAATAATAGGT----- 690
QY 230 GluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeu 249
Db |||:|||||:
691 -----TTACAAATTTCTTTGAGCAGGAGAAATCTGAG-----CCTGATGTTTCAAA 738
QY 250 GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg---MetPheIleMetThr 268
Db |||:|||||:
739 CAACAACTGTGGAATCATTTACAAAGTGTTACAAAGATTCAAACTTCATTTACATTTGACGG 798

QY 269 AsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLysGluLeuValTrpLys 288
Db |||:|||||:
799 GATGAAGATACATATTTTGGACACGCACTAAAA-----GGCAACCTGCTTGGGAA 849
QY 289 LysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArg 308
Db |||:|||||:
850 GAATTAAGATTAAAAAGAAAAAAGAAATGTATAAATTAATAAGGCGACACTTGGATTTTC 909
QY 309 SerArgGlyLeuLysSerSerGluMetAlaLysGluLeuValGluLeu 325
Db |||:|||||:
910 GGT---GGTCTGAGTCTGCAACATCTTTAGCAACACAAGTAGCAGATGTA 957
RESULT 36
ABK74770
ID ABK74770 standard; DNA; 944 BP.
XX
AC ABK74770;
XX
DT 13-AUG-2002 (first entry)
XX
DE Bacillus licheniformis genomic sequence tag (GST) #2061.
XX
KW Differential gene expression; genomic sequenced tag; GST;
KW altered culture condition; environmental stress;
KW physiological provocation; ds.
XX
OS Bacillus licheniformis.
XX
PN WO200229113-A2.
XX
PD 11-APR-2002.
XX
PF 05-OCT-2001; 2001WO-US031437.
XX
PR 06-OCT-2000; 2000US-00680598.
XX
PA 27-MAR-2001; 2001US-0279526P.
XX
PA (NOVO) NOVOZYMES BIOTECH INC.
XX (NOVO) NOVOZYMES AS.
PI Berka R, Clausen IG;
XX
XX WPI; 2002-416684/44.
XX
PT Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second Bacillus
PT cells, by using substrate containing Bacillus genomic sequenced tag
PT array.
XX
PS Claim 4; SEQ ID NO 2061; 200pp; English.
XX
XX The invention describes a method of monitoring differential expression of
XX genes in a first Bacillus cell relative to expression of the genes in
XX other Bacillus cells, comprising hybridising labelled nucleic acid probes
XX isolated from Bacillus cells to a substrate containing array of Bacillus
XX genomic sequenced tags (GST), examining the array, and determining
XX relative gene expression by an observed hybridisation reporter signal of
XX a spot in the array. The method is useful for measuring the expression of
XX genes in a first Bacillus cell relative to expression of the same genes
XX in one or more second Bacillus cells. The method is useful for monitoring new
XX global expression of several genes from a Bacillus cell, discovering changes
XX in expression of genes may be used to provide a representation of the way
XX monitoring gene copy number variation and stability. Monitoring of the way
XX in which Bacillus cells adapt to changes in culture conditions,
XX environmental stress or other physiological provocation. Extensive follow
XX up characterisation is unnecessary, when one spot on an array equals one
XX gene or one open reading frame, since sequence information is available.
XX This sequence represents a genomic sequence tag (GST) used in the method
XX of the invention. Note: The sequence data for this patent did not form
XX part of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 944 BP; 332 A; 197 C; 207 G; 208 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.6e-27	Length:	944
Score:	380.50	Matches:	109
Percent Similarity:	52.92%	Conservative:	63
Best Local Similarity:	33.54%	Mismatches:	128
Query Match:	22.51%	Indels:	25
DB:	6	Gaps:	13

US-10-724-972A-6352 (1-335) x ABK74770 (1-944)

QY	8	LeuLysGluLeuSer---ValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCys	26
DB	1	TTGAAAAAACTATCTTTATTGATTATGGCATTTGTAACCGTCTTCTGTCGTAGCGCGCTGC	60
QY	27	GlyAsnAsnSerSerSerAsnSerSerLysGluSerSer---LysAspGlyValGluIle	45
DB	61	GGAAACAAGCTGCTGCACACAGATGCANAATCTTCTCTAAAAAAGAAACCGTCACAATT	120
QY	46	LysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeuGlu	65
DB	121	AAACAACAAGCGCACACACAGATGTACGAAACCCGAAAGAAAGTCGTGTCTTCAAT	180
QY	66	TyrSerPheValAspAlaLeu-----ValAlaLeuAspValLysProValGlyIleAla	83
DB	181	TTTGGAAATGCTTGACACGCTTGATGAGCTTGGCTTAAGCGACCGCTTGTGGCCCTTCG	240
QY	84	AspAspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThr	103
DB	241	AAACAACACCTTCCGTCT---TACTTGAATAAATATATAAGAC-----AGCAAAATTGAA	291
QY	104	SerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeu	123
DB	292	AGCGTCGGCGGACTAAAGAGCCAGATTTTGAAAATAATTCGTAGCTTGAGCGCTGATTG	351
QY	124	IleIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaPro	143
DB	352	ATCATCATCGAA---GGCAGACAGCTGAACAGTATGACGAGTTTAAAAAGATCGCTCG	408
QY	144	ThrIleGluLeu-----LysSerPheAspGlyAspTyrAsnGluAsnIleAsp	159
DB	409	ACGATCGACATTGACGTTGCACAGAAAGACTACATGAACAGTTTAAAAAGAAATACCGAA	468
QY	160	AlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGluGlyLysArgLeuGlu	179
DB	469	-----AACATTGGAAAATACTTCGCGCAAGAAGAGATGCTGTCAAGAGCCCTTCGC	519
QY	180	GluHisAspLysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLys	199
DB	520	AAAAATCGATGAAAAAGTAGAGAAAGCCAAACAG---CTGGCGAAAGAAAAGCGGCGCAA	576
QY	200	ValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrVal	219
DB	577	GGATTGTGTGCTTAACAGAGCGAATAAGCAAAATCAAGCGGTATGGAGCGGCTCAAGATTC	636
QY	220	GlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAspValThrLysGly	239
DB	637	GGCTTGATCCAGATGTCTCGCGTT-----GAACTCTCGATATAAAAC	681
QY	240	LeuSerLysTyrLeuLysGlyProTyrThrLeuGlnMetAsnThrGluThrLeuSerGlnVal	259
DB	682	ATCGAGGTTTCCCTTCCAGCGCAA-----AGCGTCTCTCAGAAATATATCTCTGAAAG	735
QY	260	AsnProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluPro---SerLeu	278
DB	736	AATCCTGATTACCTGTTTGTTCATCGACCGAAGCGCTGCCATCGCGCAAAAAGAACGCA	795
QY	279	LysGluLeuGluLysAspProValTTrpLysLysLeuAsnAlaValLysAsnGlnArgVal	298
DB	796	AAAGAAGTCATCGAAAATGAATTTGAATCAACAAGGCGCTATAAAAAGCGGCATATC	855
QY	299	AspIleLeuAspArgAspLeuTrpAlaArgSer---ArgGlyLeuIleLeuSerSerGluGlu	317

DB	856	ACATACCTTGTATCCAGGCTACTTGGTACCTGTACGGCGGAGGCGTTGAATCCGTATCTGAA	915
Qy	318	MetAlaIySGLuLeu 322	
Db	916	ATGGTGAAGAAGTC 930	
RESULT 37			
AAS71379/c			
ID	AAS71379	standard; cDNA; 2209 BP.	
XX	AAS71379;		
XX			
DT	13-FEB-2002	(first entry)	
DE	DNA	encoding novel human diagnostic protein #7183.	
XX			
KW	Human;	chromosome mapping; gene mapping; gene therapy; forensic;	
KW	food supplement;	medical imaging; diagnostic; genetic disorder; ss.	
XX			
OS	Homo sapiens.		
XX			
PN	WO200175067-A2.		
XX			
PD	11-OCT-2001.		
XX			
PF	30-MAR-2001;	2001WO-US008631.	
XX			
PR	31-MAR-2000;	2000US-00540217.	
PR	23-AUG-2000;	2000US-00649167.	
XX			
PA	(HYSE-)	HYSEQ INC.	

XX

PI Drmanac RT, Liu C, Tang YT;

[illegible]

DR WPI; 2001-639362/73.
DR D-PCNB: ADG07193

XX
F F00B, ABG0/132.

PT New isolated polynuc

PT diagnostics, forensi

PT responsible for gene

PT biodiversity.

ax
PS
Claim 1: SEO TD NO 7

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

CC The invention relates

sequences. (I) is us

CC reaction (PCR) prime

and in recombinant p
in diagnostics as ex

CC
...
genes. (I) is useful

CC activity of (II) or

CC useful for generating

polypeptide in tissu

CC supplement. (II) and
CC of gitaas representing

of genes expressing
involving aberrant n

CC polypeptide and poly

CC diagnostics, forensi

responsible for gene

CC and to produce other

amino acid sequences
coding sequences of

CC
CC
CC

CC electronic format di:

ftp.wipo.int/pub/pub:

XX

sequence Z209 BP; 480

Alignment Scores:

Pred. No.: 2

Score: 3

Percent Similarity: 50

100

2

XX	DNA encoding novel human diagnostic protein #8842.
DE	Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX	food supplement; medical imaging; diagnostic; genetic disorder; ss.
KW	Homo sapiens.
XX	WO200175067-A2.
XX	11-OCT-2001.
PD	30-MAR-2001; 2001WO-US008631.
XX	31-MAR-2000; 2000US-00540217.
XX	23-AUG-2000; 2000US-00649167.
XX	(HYSB-) HYSEQ INC.
PA	Drmanac RT, Liu C, Tang YT;
XX	P-PSDB; ABG08851.
DR	New isolated polynucleotide and encoded polypeptides, useful in
XX	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.
XX	Claim 1; SEQ ID NO 8842; 103pp; English.
PS	The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX	sequences. (I) is useful as hybridisation probes, polymerase chain
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC	and in recombinant production of (II). The polynucleotides are also used
CC	in diagnostics as expressed sequence tags for identifying expressed
CC	genes. (I) is useful in gene therapy techniques to restore normal
CC	activity of (II) or to treat disease states involving (II). (II) is
CC	useful for generating antibodies against it, detecting or quantitating a
CC	polypeptide in tissue, as molecular weight markers and as a food
CC	supplement. (II) and its binding partners are useful in medical imaging
CC	of sites expressing (II). (I) and (II) are useful for treating disorders
CC	involving aberrant protein expression or biological activity. The
CC	polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC	coding sequences of the invention. Note: The sequence data for this
CC	patent did not appear in the printed specification, but was obtained in
CC	electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	Sequence 4392 BP; 1063 A; 1157 C; 1181 G; 990 T; 0 U; 1 Other;
SQ	
Alignment Scores:	
Pred. No.:	6.05e-26
Score:	375.00
Percent Similarity:	50.17%
Best Local Similarity:	32.54%
Query Match:	22.19%
DB:	5
US-10-724-972A-6352 (1-335) x AAS73038 (1-4392)	
QY	45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64
DB	70 GTTCAGGACGACACGCGCGATTAACTCGAAACACGCGACCGATTGTGGTCTG 129
QY	65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84
DB	130 GAACTCTCGTTCCCGGATCGCTGGCGCGGTGACGCTCATCCGATCGGTATTGCCGAC 189

XX	Best Local Similarity: 32.54%	Mismatches: 115
XX	Query Match: 22.19%	Indels: 32
XX	DB: 5	Gaps: 5
US-10-724-972A-6352 (1-335) x AAS71379 (1-2209)		
QY	45	IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64
DB	2140	GTTTCAGGACGACACGCGCGATTAACTCGAAACACGCGACCGATTGTGGTCTG 2081
QY	65	GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84
DB	2080	GAACTCTCGTTCCCGGATCGCTGGCGCGGTGACGCTCATCCGATCGGTATTGCCGAC 2021
QY	85	AspAsnLysLysAsnArgLleLleLysProLeuArgAspLysLleGlyLysTyrThrSer 104
DB	2020	GATAACGATCAAAACGATCTCTCCCGAAGTCGTCGACCTGAAACCGTCGACGATCC 1961
QY	105	ValGlyThrArgLysGlnProAsnLeuGluGluLysSerLysLeuLysProAspLeuIle 124
DB	1960	GTCGGAACGCGCGACGCGCTGGAAGCCATTGCCGCTCTGAAACACGACCTGATC 1901
QY	125	IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysLleAlaProThr 144
DB	1900	ATTGCCGACGACGCTGCCATCGGGGGTTTACATCGCTTGACGAAATCGCGCGGTA 1841
QY	145	IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnLleAspAlaPheLysThrIle 164
DB	1840	CTGCTGCTTAAGTCCCGCAACGAAACCTACGCTGAAATTTGCAATCTCGCGGTATCATC 1781
QY	165	SerLysAlaLeuGlyLysGluGluGluGlyLysArgLeuGluGluHisAspLysLys 184
DB	1780	GGCGAAATGGTGGGTAAACGAGAGATGACGACGCTCTGGAACACATAAGAGAGG 1721
QY	185	IleGluGluTyrLysLysGlyIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204
DB	1720	ATGGCGCAGTGGCGGACGAG-----CTTCCCAAGGACACGCGTGGCTTTGGCACA 1667
QY	205	AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224
DB	1666	TCACGGAACAGCAATTCAACCTGCATCTACGAGACCTGGACCGGCGGTGCTGGCC 1607
QY	225	GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu 244
DB	1606	TCTCTGGGGCTG-----AACGTTCCCGCTCGATG 1577
QY	245	LysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMet 264
DB	1576	GCGGTGCTGCCATCCGCTCGCTGCTGGCGGTTCATGTCCACCGCGATGT 1517
QY	265	PheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGluLysAsp 284
DB	1516	TCTATCCATGCAATTACGTTTACATCAACACACACCTGTCTC----- 1475
QY	285	ProValTyrLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeu----- 301
DB	1474	-----TGGACCGCGGGTCTGTGATCCGCTGCGGTGGGTCTGAAATGACC 1421
QY	302	-----AspArgAspLeuTyrPalaArgSerArgGlyLeuLysSerSerGlu 316
DB	1420	GACGAAGCGGTGAAGATCGGAACCTGGTTCCGCTTCCGACACGCAAGCTGACGAAGAA 1361
QY	317	-----GluMetAlaLysGluLeuValGluLeuSerLys 327
DB	1360	TACGATCACATTAAAGACGTTAACGATCTGCTGCACTCTGAACTGCTGAAA 1316
RESULT 38		
AAS73038		
ID	AAS73038 standard; cDNA; 4392 BP.	
XX		
AC	AAS73038;	
XX		
DT	13-FEB-2002 (first entry)	

Db 242 TTGAATGACGTTAAAAATATCGGTGATATTGTTGACCGTGATGTTGGCAAGATCT 301
QY 310 ArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLeuAsp 329
Db 302 CGTGGCTTAATTTCTTCGAGAAATGCTAAAGAACTTGTGAATTATCAAAAAAGAA 361
QY 330 SerLys 331
Db 362 CAAAAG 367
RESULT 40
ID AAS77335
AC AAS77335 standard; cDNA; 2799 BP.
XX AAS77335;
DT 13-FEB-2002 (first entry)
DE DNA encoding novel human diagnostic protein #13139.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Dmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG13148.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
XX biodiversity.
PS Claim 1; SEQ ID NO 13139; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 2799 BP; 619 A; 740 C; 776 G; 664 T; 0 U; 0 Other;
SQ

Alignment Scores:

Pred. No.: 9,06e-24 Length: 2799
Score: 350.00 Matches: 69
Percent Similarity: 68.92% Conservative: 46
Best Local Similarity: 46.62% Mismatches: 33
Query Match: 20.71% Indels: 0
DB: Gaps: 0
US-10-724-972a-6352 (1-335) x AAS77335 (1-2799)
QY 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64
Db 1672 GTTCAGACGACACCGCAGCTTACACTCGAAAAACGCCACACGATTTGTGGTCTG 1731
QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84
Db 1732 GAACTCTCGTTCCGCGATGCGCTGGCCGCGTCCGATCATCCGATTCGGTATTGCCGAC 1791
QY 85 AspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSer 104
Db 1792 GATAACCGATCAAAACGCATCCTGCCCGAAGTGGTCCGCCACCTGAAACCGTGGCAGTCC 1851
QY 105 ValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLysLeuLysProAspLeuIle 124
Db 1852 GTCGAAACGCGCGCGAGCGAGCTGGAAGCCATTGCCCGCTCTGAAACCGACCTGATC 1911
QY 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144
Db 1912 ATTGCCGACAGCAGTCGCCATGCGGGGTTTACATCGCTTGACAGCAATTCGCCCGGTA 1971
QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164
Db 1972 CTGCTGCTTAAGTCCCGCAACCAACCTACGCTGAAAATTTGCAATCTCGCGGCTATCATC 2031
QY 165 SerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAspLysLys 184
Db 2032 GCGCAATGTTGGGTAAAGCGGAGAGATGCGACGACGTCTGGAAACAACATAAAGAGAGG 2091
QY 185 IleGluGluTyrLysLysGluIle 192
Db 2092 ATGGCGCAGTGGCGCAGCGAGCTT 2115

Search completed: November 9, 2005, 19:55:15
Job time : 749 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: November 9, 2005, 19:55:21 ; Search time 3860 Seconds
(without alignments)
3303.505 Million cell updates/sec

Title: US-10-724-972A-6352
Perfect score: 1690
Sequence: 1 GVESVGLKILSVIGLLFVL.....EEMAKELVLSKKDKKONK 335

Scoring table:
BLOSUM62 Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-O=/cgn2.1/USPTO.spool_h/US10724972/runat_03112005_141900_29152/app_query.fasta_1.519
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=DISSIM2 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10724972 @CNG 1.1 3437 @runat_03112005_141900_29152 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBIOCK=100 -LONGLOG
-DSV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gsal:
9: gb_gsal2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	412	24.4	776	9	CL694436
c 2	410	24.3	801	9	CL674627
c 3	368	21.8	625	7	CF307287
c 4	210	12.4	531	2	BE432885
c 5	152	9.0	760	9	CL682049
c 6	129	7.6	966	9	CNS01MUR
c 7	126.5	7.5	2019	9	CL958706
c 8	122	7.2	810	9	CL658559
c 9	122	7.2	881	9	CL666398

10	121.5	7.2	4137	9	AY420057
c 11	119.5	7.1	764	4	BJ347654
c 12	119	7.0	839	8	AZ532836
c 13	119	7.0	906	8	AZ538305
c 14	119	7.0	1782	8	CL957361
c 15	118.5	7.0	897	8	BH132864
c 16	118.5	7.0	2740	3	AK029960
c 17	117.5	7.0	1005	5	BQ065769
c 18	117	6.9	1065	7	CF109973
c 19	117	6.9	5504	9	AY407955
c 20	116.5	6.9	804	7	CO573945
c 21	116.5	6.9	832	4	BJ714899
c 22	116.5	6.9	1545	3	AK012454
c 23	116	6.9	816	7	CO113362
c 24	116	6.9	889	8	AZ549761
c 25	116	6.9	1483	3	CR650378
c 26	115.5	6.8	673	1	AL854856
c 27	115.5	6.8	684	5	BW253010
c 28	115.5	6.8	690	5	BW254793
c 29	115.5	6.8	692	5	BW222839
c 30	115.5	6.8	700	5	BW235942
c 31	115.5	6.8	702	5	BW248137
c 32	115.5	6.8	713	5	BW441129
c 33	115.5	6.8	753	5	BW055754
c 34	115.5	6.8	981	4	BI905625
c 35	115.5	6.8	4107	9	AY420059
c 36	115	6.8	785	7	CR426562
c 37	115	6.8	1120	3	CR731171
c 38	115	6.8	2252	3	CR685229
c 39	114.5	6.8	707	4	BJ706325
c 40	114.5	6.8	711	1	AL666132
c 41	114.5	6.8	723	4	BJ714400
c 42	114.5	6.8	727	4	BJ712865
c 43	114.5	6.8	733	4	BJ707678
c 44	114.5	6.8	751	4	BJ712289
c 45	114.5	6.8	811	4	BJ714906

ALIGNMENTS

RESULT 1
CL694436/c
LOCUS
DEFINITION
CL694436
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
CL694436 776 bp DNA linear GSS 10-JUL-2004
PRI0164b_A02.2 - PRI0164b.BR (776) Mixed stage fosmid library of P.
pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.
CL694436.1 GI:50216344
GSS.
Pristionchus pacificus
Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
1 (bases 1 to 776)
Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
AppADB: an AcedB database for the nematode satellite organism
Pristionchus pacificus
Nucleic Acids Res. 32 (1), D421-D422 (2004)
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end
sequenced at Vancouver, Canada.
Seq primer: 17
Class: fosmid ends.
Location/Qualifiers
1..776
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"


```

QY 185 IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204
   ::::::::::: ::::: ::::: ::::: ::::: :::::
DB 499 ATGGCCAGTGGCCAGCCAG-----CTTCCCAAAGGACACGCGTGGCTTTGGCACA 552

QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224
   ::::: ::::: ::::: ::::: ::::: :::::
DB 553 TCACGGCAACAGCAATTAACCTGCATACTCAGGACACTGGACCGGAGCGTGTGGCC 612

QY 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu 244
   ||||| ::::: ::::: ::::: ::::: :::::
DB 613 TCTCTGGGGCTG-----AACGTTCCCGCTGGGATG 642

QY 245 LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263
   ||| ::::: ::::: ::::: ::::: :::::
DB 643 GCGGGTGGCTGCCATCGCGCTGGACAACTGCTGGCGGTCAATCTGCTGCTGG 702

QY 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGlyLys 283
   ::::: ::::: ::::: ::::: :::::
DB 703 CTGCTGGTGGCCACTATCGC-----GAAGAGAGCATGTTAAACGCTGGCAACAA 753

QY 284 AspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgVal 298
   ||||| ::::: ::::: ::::: :::::
DB 754 GATCCGCTCTGGCAGATGTTAAACGCGCGCAGAACACGAGTT 798

RESULT 3
CF307287/c
LOCUS
DEFINITION
  HDAL--06-E11.g1 OshDAC1-overexpressing transgenic rice lambda phage
  cDNA library I (HDAL) Oryza sativa (japonica cultivar-group) cDNA
  clone HDAL--06-E11, mRNA sequence.
CF307287
CF307287.1 GI:33679048
EST.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 625)
Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
Large-scale Sequencing Analysis of Rice ESTs
Unpublished (2003)
Contact: Nahm B.H.
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
of Bioscience and Bioinformatics, Myongji University
Yongin, Kyeonggi, Korea
Tel: 82 31 320 6193
Fax: 82 31 321 6355
Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.
FEATURES
    Location/Qualifiers
        1..625
            /organism="Oryza sativa (japonica cultivar-group)"
            /mol_type="mRNA"
            /cultur="Nackdong"
            /db_xref="taxon:39947"
            /clone="HDAL--06-E11"
            /tissue_type="callus"
            /dev_stage="proliferated callus on 2N6 media for 2 weeks"
            /lab_host="E.coli SOLR"
            /clone_lib="OshDAC1-overexpressing transgenic rice lambda
            phage cDNA library I (HDAL)"
            /notes="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2:
            XhoI; Callus was treated with ABA(20um) for 1 hour. cDNA
            was inserted into lambda Uni-ZAP XR vector at 5' end with
            EcoRI and 3' end with XhoI site. mRNA was derived from
            rice Histone Deacetylase overexpression line."
ORIGIN
Alignment Scores: 9.15e-31 Length: 625
Pred. No.: 369.00 Matches: 76
Score:

```

```

Percent Similarity: 62.84% Conservative: 39
Best Local Similarity: 41.53% Mismatches: 66
Query Match: 21.78% Indels: 2
DB: 7 Gaps: 1

US-10-724-972A-6352 (1-335) x CF307287 (1-625)

QY 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64
   ::::::::::: ::::: ::::: ::::: ::::: :::::
DB 619 GTTCAGGACGACACGCGCAGTTTACACTCGAAAAAAGCCCAACGATTTGTGTGTTG 560

QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84
   ||||| ::::: ::::: ::::: ::::: :::::
DB 559 GAACTCTCTGTTCCGATGCGCTGGCCGCGTGGACGTCATCCCATCGTATTGGCGAC 500

QY 85 AspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSer 104
   ||||| ::::: ::::: ::::: ::::: :::::
DB 499 GATAACGATGCAAAACGCGCATCTGCGGAGTGGCTGGCGACCTGAAACCGTGGCAGTCC 440

QY 105 ValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIle 124
   ||||| ::::: ::::: ::::: ::::: :::::
DB 439 GTTCGAACGCGCGCGCAGCGCGAGCTTGAAGCCATTGCGCTCTGAAACCGACACCTGATC 380

QY 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144
   ||||| ::::: ::::: ::::: ::::: :::::
DB 379 ATTGCGGACGACGATCGCGGCGTGTACATCGCTTGCACAAATCGCGCGGTA 320

QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164
   ::::: ::::: ::::: ::::: :::::
DB 319 CTGCTGCTTAAGTCCCGCAACGAAACCTACGCTGAAATTTGCAATCTCGCGCTATCATC 260

QY 165 SerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAspLysLys 184
   ::::: ::::: ::::: ::::: :::::
DB 259 GCGCAATGCTGGGTAAAAAGCAGAGATGACGAGCAGCTCTGGAACAACATAAGAGAGG 200

QY 195 IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204
   ::::: ::::: ::::: ::::: :::::
DB 199 ATGGCGCAGTGGCGCGCAGCAG-----CTTCCAAAGGACACGCGTGGCTTTGGCACA 146

QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224
   ::::: ::::: ::::: ::::: :::::
DB 145 TCACCGCAACAGCAATTAACCTGCATACTCAGGAGACCTGGACCGGACGCTGTGGCC 86

QY 225 GlnLeuGly 227
   |||||
DB 85 TCTCTGGGG 77

RESULT 4
BE432885
LOCUS
DEFINITION
  BE432885 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
  clone cLEG10K14, mRNA sequence.
ACCESSION
  BE432885
VERSION
  BE432885.1 GI:9430644
SOURCE
  Lycopersicon esculentum (tomato)
  Lycopersicon esculentum
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
  1 (bases 1 to 531)
  Liang,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,T.E.,
  Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Renning,C.M.,
  Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and
  Tanksley,S.D.
  Generation of ESTs from tomato fruit tissue, breaker stage
  Unpublished (2000)
  Contact: CUGI
  Clemson University Genomics Institute
  Clemson University
  100 Jordan Hall, Clemson, SC 29634, USA
  Email: http://www.genome.clemson.edu/orders/index.html
  5 prime sequence.

```

FEATURES
source

Location/Qualifiers
1. .531
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CLEG10K14"
/tissue_type="Pericarp"
/dev_stage="breaker"
/lab_host="SOLa"
/clone_lib="tomato breaker fruit, TIGR"
/note="Vector: pBluescriptSKmCuadapt; Site 1: EcoRI;
Site 2: XhoI; Fruit were harvested at the breaker stage
(first sign of lycopene accumulation on the blossom end of
the fruit). Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."

ORIGIN

Alignment Scores:
Pred. No.: 4.28e-13 Length: 531
Score: 210.00 Matches: 41
Percent Similarity: 75.3% Conservativeness: 11
Best Local Similarity: 59.4% Mismatches: 17
Query Match: 12.43% Indels: 0
DB: 2 Gaps: 0

US-10-724-972A-6352 (1-335) x BE432885 (1-531)

QY 67 SerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAspAsn 86
|||||
Db 4 TCGTTCCGCGATGCGTGGCCCGTGGACGTCATCCGATCGGTATTGCCGACGATAAC 63
QY 87 LysLysAsnArgIleLeuLysProLeuArgAspLysIleGlyLysThrSerValGly 106
|||||
Db 64 GATGCAAAACGATCTCTGCCGAAGTGGTGGCCACCTGAACCGTGGCAGTCCGTCGA 123
QY 107 ThrArgLysGlnProAsnLeuGluLysSerLysLeuLysProAspLeuIleAla 126
|||||
Db 124 ACGCGCGCGACGCGAGCTGGAAGCATTTGCCGCTCTGAACACGACCTGATTCGC 183
QY 127 AspAsnAsnArgHisLysGlyTleTyr 135
|||||
Db 184 GACAGCAGTGCCTATCGCGGGGTTTAC 210

RESULT 5

CL682049 760 bp DNA linear GSS 09-JUL-2004
LOCUS PRI0132d_D03.2 - PRI0132d.BR (760) Mixed stage foamid library of P.
DEFINITION pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.

CL682049 GI:50189271
CL682049.1

GSS.

Pristionchus pacificus

Pristionchus pacificus

Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;

Neodiplogasteridae; Pristionchus.

1 (bases 1 to 760)

Srinivasan, J., Otto, G.W., Kahlow, U., Geisler, R. and Sommer, R.J.

AppDB: an AcedB database for the nematode satellite organism

Pristionchus pacificus

Nucleic Acids Res. 32 (1), D421-D422 (2004)

Contact: Sommer RJ

Evolutionary Biology

Max-Planck-Institute for Developmental Biology

Spemannstr. 37-39, Tuebingen D-72076, Germany

Tel: 00497071601371

Fax: 00497071601498

Email: ralf.sommer@tuebingen.mpg.de

This library was generated at Caltech, Pasadena, USA and end

sequenced at Vancouver, Canada.

Seq primer: T7

Class: foamid ends.

FEATURES
source

Location/Qualifiers
1. .760
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage foamid library of P. pacificus
var. California"
/note="Vector: pEpifos-5 Foamid vector"

ORIGIN

Alignment Scores:
Pred. No.: 2.36e-06 Length: 760
Score: 152.00 Matches: 62
Percent Similarity: 42.91% Conservativeness: 50
Best Local Similarity: 23.75% Mismatches: 111
Query Match: 8.99% Indels: 38
DB: 9 Gaps: 11

US-10-724-972A-6352 (1-335) x CL682049 (1-760)

QY 60 ArgValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysPro 79
|||||
Db 3 CGTATTGTTTCCACCAGCGTCACCGCTGACCGGCTCACTGTGGCGATTGATGCTCCGGTG 62
QY 80 ValGlyIleAlaAspAsnLysLysAsnArgIleLysPro-----LeuArg 96
|||||
Db 63 ATCGCCAGCGCGCGACCAACCGGATTAACCGGTGCGGATGACCGGGCTTTTACGC 122
QY 97 -----AspLysIleGlyLysTyrThrSerValGlyThrArgLys----- 109
|||||
Db 123 CAGTGGAGCAAGTGGCGAAA-----GAACGCAAACTGCAACGGCTCTAT 167
QY 110 -----GlnProAsnLeuGluLysSerLysLysProAspLeuIle----- 125
|||||
Db 168 ATCGCGCAACGAGCGCGGCGGAGCGGTTGCCGCAAAATGCGCGATCTGATTTAATTAGC 227
QY 126 --AlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144
|||||
Db 228 GCAACCGCGGGGATTCGGCGCTGGCAGTGTATCATGCTTTCATATCGCCCGACA 287
QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164
|||||
Db 288 TTAATCATCAATTACGACGACAAAAAGCTGGCAGTCGCTTTAACGCGAGTTGGCGAAATT 347
QY 165 SerLysAlaLeuGlyLysGluGluGlyLysArgLeuGluGluHiAspLysLys 184
|||||
Db 348 ACC-----GGCATGAGAAAACAAGCGCAGCGGATTCACAGTTTGTGATAAGCAA 398
QY 185 IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204
|||||
Db 399 CTGGCAGCGCGGAAGAGCAATCAATTACCGCGCGAG-----CCGCTCACT 446
QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsn-----SerTyrVal 219
|||||
Db 447 GCCATTGCTATACGCGCGTGCACAGTCCCAATCTCTGCAGCGCGGAATTCAGACAA 506
QY 220 GlyGlnPheLeuSerGlnLeuGlyPheLysGluAla---LeuSerAspAspValThrLys 238
|||||
Db 507 GGGCAGATGCTGGAAACAACCTCGGCTTTCAGCTGGCGAAGTTGCCCGCAGGGTTAAACGCC 566
QY 239 GlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGln 258
|||||
Db 567 AGCCAAAGCCAGGCGCAACGCATGACATCAATTGCTGTGGGGAATAATCTGGCTCGG 626
QY 259 ---ValAsnProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSer 277
|||||
Db 627 GGGTTAAACGGTGAATCGCTATTCTCTGTTTC-----GCCGGTGTATCAGAAAGAC 674
QY 278 LeuLysGluLeuGluLysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArg 297
|||||
Db 675 GCCGATGCGATTATATGCCAATCCACTGCTCGCGCACCTGCTCGCGCAGTACAAACAGCAG 734

```
QY 298 Val 298
Db 735 GTT 737

RESULT 6
LOCUS CNS01MUR 966 bp DNA linear GSS 14-JUN-2001
DEFINITION Anopheles gambiae GSS T7 end of clone 22J04 of NotreDamel library
from strain PEST of Anopheles gambiae (African malaria mosquito),
genomic survey sequence.
ACCESSION AL151412
VERSION AL151412.1 GI:7011891
KEYWORDS GSS.
SOURCE Anopheles gambiae (African malaria mosquito)
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Anopheles.
REFERENCE 1 (bases 1 to 966)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

REFERENCE 2 (bases 1 to 966)
AUTHORS Roth,C.W., Brey,P.T., Ke,Z., Collins,F.H. and Weissenbach,J.
TITLE Direct Submission
JOURNAL Submitted (16-FEB-2000) BIMI, Institut Pasteur, 25, rue du Dr.
Roux, Paris 75015, France
COMMENT This clone is from an A. gambiae BAC library provided by F.H.
Collins and sequenced by Genoscope in collaboration with the
Laboratory of Biochem. and Biol. Molec. of Insects, Institut
Pasteur.

FEATURES
source
1..966
Location/Qualifiers
/organism="Anopheles gambiae"
/mol_type="genomic DNA"
/strain="PEST"
/db_xref="taxon:7165"
/clone="22J04"
/clone_lib="NotreDamel"
/note="end : T7"

ORIGIN
Alignment Scores:
Pred. No.: 0.00128 Length: 966
Score: 129.00 Matches: 67
Percent Similarity: 43.21% Conservative: 54
Best Local Similarity: 23.93% Mismatches: 116
Query Match: 7.63% Indels: 44
DB: 9 Gaps: 11

US-10-724-972a-6352 (1-335) x CNS01MUR (1-966)

QY 44 GluileYHieGluglyThrLysValProLYsHieProLYsHieProLYsValValVal 63
Db 160 CAGTGACCCAGCAGCGCGGGTACATCCCTTGAGCACAACCCGACGGCATGCTCC 219
QY 64 LeuGluTYrSerPheValaspAlaLeuValAlaLeuAspVallyProValGly-IleAl 83
Db 220 ACCAGCGTGACCTTAACCGGCTCCCTGCTGGCCATTRACGCACCGGCTGCCAGCGGS 279
QY 83 aAspAspAsnLYsAsnArg- - - - -IlelleLYsProLeuArgAspLY 98
Db 280 CGACCAAGCGGAGAACACCGGCTGGCGNATGGACAGGGTTCTCTGCCCA- - - - - 328
QY 98 sLleGlyLYsTYrThrSerValGlyThrArgLYs- - - - -GlnPr 111
Db 329 -GTGGGGCGCATATGGCGAA- - - - -CARCGAAGTTGCCGCTGTATATCGCGAGCC 381
QY 111 oAsnLeuGluGluIleSerLYsLeuLYsProAspLeuIlelleAlaAspAsnArgHi 131
Db 382 GAGCGCTGAAGCGGTGGCGCGCAAAATTCGGACCTKATTTTTTAT-CAGCGCCACCGSG 440
```

```
QY 131 sLYs- - - - -GlyIleTYrLYsAspLeuAsnLYsIleAlaProThrIleGluLeuLY 148
Db 441 GGATTCCCGCGTGGCGCTGTACGATCAGCTTTCCGCCATTGCGCGAGCGTTATCATCAA 500
QY 148 sSerPheAspGlyAspTYrAsnGluAsnIleAspAlaPheLYsThrLYsLYsAlaLe 168
Db 501 CTACKACCAAGAGAGCTGGCAGGTGCTCTGACCCAGCTGGCGGAGATCACC- - - - - 552
QY 168 uGlyLYsGluGluGlyLYsLYsArgLeuGluGluHisAspLYsLYsIleGluGluTY 188
Db 553 -GGACACGAAAGCAGCGCGCAGCGCATTTGCGCGTTTGATAAACAGCTCGCGCAGGT 611
QY 188 rLYsLYsGluIleThrMetAspLYs- - - - -AsnGlnLYsValLeuProAlaValAl 205
Db 612 GAAACAGCAGATGACGCTGCCCGCCGACGCCGCGTGAACCCATCTGCTACACCGCGCGCG 671
QY 205 aAlaLYsSerGlyLeuLeuAlaHisProSerAsnSerTYrValGlyGlnPheLeuSerGl 225
Db 672 GCATTCCCGCCAACTTTTGACC- - - - -ACGGAATCGGCACAGGCAAGCTCTGCACCA 725
QY 225 nLeuGlyPheLYsGluAlaLeuSerAspAspValThrLYsLYsGlyLeu- - - - -Se 241
Db 726 GCTGGGCTTTAGCGTGGCG- - - - -GACCTGCTGCGCGGTTGCAGACCTCAACAAG 776
QY 241 rLYsTYrLeuLYsGlyProTYrLeuGlnMetAsnThrGluThrLeuSerGln- - ValAsn 260
Db 777 CCAGGCGCAAGCGCCACGACATTTATCCAGCTGGCGCGTGAACCTKGCAGCGGGCTGAAC 836
QY 261 ProGluArgMetPheIleMetThrAsnLYsAlaSerSerAsnGluProSerLeuLYsGlu 280
Db 837 GCGCAAGCGGTGTTTGTGTTT- - - - -GCCGCTGACGAGAAAGACCTKGCGGCG 884
QY 281 LeuGluLYsAspProValTriPlyLYsLYsLeuAsnAlaValLYsAsnGlnArgVal 298
Db 885 ATTTCAGCTAACCCGCTGCTKGCACCTGCGCTGGTAAAGAAAACAAGCGCGTA 938

RESULT 7
LOCUS CL958706 2019 bp DNA linear GSS 21-SEP-2004
DEFINITION OaIFCC001947 Oryza sativa Express Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.
ACCESSION CL958706
VERSION CL958706.1 GI:52372164
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 2019)
AUTHORS Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
Wong,G.K.S., Deng,X.W. and Wang,J.
TITLE An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
JOURNAL Unpublished (2004)
COMMENT Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
Location/Qualifiers
1..2019
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences "
```



```
Db 3076 AAAACAGACGAGCTAAACAAACAGTTAAAGAGACTTGTACAGAAATACACGGAAGTAAAG 3135
Qy 107 ThrArgLysGlnProAsnLeuGluGluLeuSerLysLeuLysProAspLeuLeuLeuLeu 126
Db 3136 AATGTGAAGAGAGAGCTAGTAGAAGAAATGCGAAACAGACTTCTGAGATACCTCGACGTG 3195
Qy 127 AspAsn-----AsnArgHis-----LysGlyIleTyrLysAsp 137
Db 3196 CAAATCTTTGCAAAACACATGTTCCATTGGACAGGTTGAGGCTCTGAAAAATCT 3255
Qy 138 LeuAsnLysIleAlaProThrIle-----GluLeuLysSerPheAspGlyAspTyrAsn 155
Db 3256 CTTAATGGCACAATTTGAAATCTAAAGGAAGAACTGAAGAGTATGCAAGGTGTACGAG 3315
Qy 156 GluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGluGlyLys 175
Db 3316 AAAGAGCAGCAGACAGTACCAAACTGCATCAATTTGAGAAATCAAAAGAACTCTTCT 3375
Qy 176 LysArgLeuGluGluHisAspLysLysIleGluTyrLysLysGluLeuLeuMet--- 194
Db 3376 GTACCCCTGGCAGACATTTGCAGATTAAAGAGCATTTGAGAAAGAGTTGGAATCATA 3435
Qy 195 -----AspLysAsnGlnLysValLeuProAlaValAlaLysSerGly 209
Db 3436 AAAGCCAGCTTCAGAGAAAGGAAGAAAGAACCCAAACAAATGCAAGAACTCCAAA 3495
Qy 210 LeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLys 229
Db 3496 CTTCTAGTCGGAGGTTTCAGAAATACT-----AAACAAGCATTTAAAAAATTAGAGACTGA 3549
Qy 230 GluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLysLeuLysGlyProTyrLeu 249
Db 3550 GAGGTAGTTGAC-----TTGTCTTAATAT---AAAGCAACAAAAGT 3588
Qy 250 GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn 269
Db 3589 GATTTGGAGACAG---ATTCTAGCTTAAT---GAAAAATTGGCCAATCTGAATAGA 3642
Qy 270 Lys-----AlaSerSerAsnGluProSerLeuLys 279
Db 3643 AAGTATGAGGAAGTATGTGAGGAAGTTTGCATGCCAAAAGAAAGAAATATCTGCAAAA 3702
Qy 280 GluLeuGlyAspProValTrpLysLysLeu---AsnAlaValLysAsnGln----- 296
Db 3703 GAT---GAGAGGAATTACTGCTATTCAGCATTTGACGAGAAATTAAGATCAGAGGNA 3759
Qy 297 ArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArgGlyLeuLeuSerSerGlu 316
Db 3760 CGATGTGAT-----AAGTCCTTAACAACAATCACA 3789
Qy 317 GluMetAlaLysGluLeuValGluLeuSerLysAspSerLysLysAspAsnLys 335
Db 3790 GAGTTTACAAAGAAAGAAATCAAGAATCTGTCTAAACAATAAGACGAAAGAAATAAAG 3846
```

```
RESULT 11
LOCUS BJ347654/1 764 bp mRNA linear EST 06-MAR-2002
DEFINITION BJ347654 Dictyostelium discoideum cDNA library, AF Dictyostelium
ACCESSION BJ347654
VERSION BJ347654.1 GI:19218161
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.
AUTHORS 1 (bases 1 to 764)
TITLES Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i.T.
Full length cDNA of Dictyostelium discoideum at the aggregation
stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadaeu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
```

```
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1. 764
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="dda27b21"
/sex="mat A"
/dev stage="Aggregation stage"
/clone_lib="Dictyostelium discoideum cDNA library, AF"

FEATURES
source
1. 764

ORIGIN
Alignment Scores:
Pred. No.: 0.0108 Length: 764
Score: 119.50 Matches: 48
Percent Similarity: 45.37% Conservative: 45
Best Local Similarity: 23.41% Mismatches: 77
Query Match: 7.07% Indels: 35
DB: 4 Gaps: 6

US-10-724-972A-6352 (1-335) x BJ347654 (1-764)

Qy 22 AlathrAlaAlaCysGlyAsnAsnSerSerAsnSerSerLysGluSerSerLysAsp 41
Db 748 TCATCATCATCATCATCATCATCATCATCATCATCATCATCTTTGGAAACAAAAAATGTA 689
Qy 42 GlyValGluIleLysHisGluGluGlyThrThrLysValProLysHisProLysArgVal 61
Db 688 GAATTTGGAAATCAACACCAACAAAATTTGGAAAGGAAATACCAAGAAATTTGGAAGGAAATAC 629
Qy 62 ValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGly 81
Db 628 CAAGAATTTGGAAGTAAATATACCAAGAA-----TTGGAAGTCTGATAACAAAGAAATTTGGA 575
Qy 82 -----ile 82
Db 574 TGTAAAAACCTAGAAATTTAAAAATCAAAACCAAGAAATTTGGAAATTAATAACAAAGAAATTTG 515
Qy 83 AlaAspAspAsnLysLysAsnArgIle---IleLysProLeuArgAspLysIleGlyLys 101
Db 514 GAAATCAAAACCAAGAAATTTGAAATCAAAATCCAAAGAAATTTAGAAACCATTAATCAATCAA 455
Qy 102 TyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLysPro 121
Db 454 TACAAAATGAAATAGTGTCCCAAGATATGGAATAATTTGATATTAATAATAAAGAAAGAA 395
Qy 122 AspLeuIleIleAlaAspAsnAsnArgHisLysGlyIle-----TyrLysAspLeu 138
Db 394 GAACCTTGTAAGAGCCATAGAGAATGAGCTAGATGATTAATAAAGGATATGTGGAACTT 335
Qy 139 AsnLysIleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIle 158
Db 334 GAAAGGTCAAT-----AAAGAAATCAAAAGAAATTTATGAAAAATCAAT 290
Qy 159 AspAlaPhe-----LysThrIleSerLysAlaLeuGlyLysGluGluGlu 173
Db 289 TTTGATTCATGAAGCAAGAAATTTGAAGAAATCAATAAATTAGTAATAATAATAATGAT 230
Qy 174 GlyLysLysArgLeuGluGluHisAspLysLysIleGluGluTyrLysLysGluLeuThr 193
Db 229 TCACAAAACCAATTTGGGAGAGAAATAATGAATGTGTACCTTTTAAAAAAGAAAGATTGAA 170
Qy 194 MetAspLysAsnGln 198
Db 169 ACATTAATAAATGAA 155

RESULT 12
LOCUS AZ532836 839 bp DNA linear GSS 03-NOV-2000
```

```

DEFINITION  ENTDA85TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
GENOMIC, genomic survey sequence.
ACCESSION   AZ532836
VERSION     AZ532836.1 GI:11088430
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica
ORGANISM    Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1 (bases 1 to 839)
AUTHORS     Loftus,B., Van Aken,S. and Fraser,C.
TITLE       Determination of clone end sequences from Entamoeba histolytica
JOURNAL     HM1:IMSS sheared DNA library
COMMENT     Unpublished (2000)
            Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjoftus@igr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: shotgun
            High quality sequence start: 21
            High quality sequence stop: 820.
            Location/Qualifiers
            1..839
               /organism="Entamoeba histolytica"
               /mol_type="genomic DNA"
               /strain="HM1:IMSS"
               /db_xref="taxon:5759"
               /clone_lib="Entamoeba histolytica Sheared DNA"
               /notes="Vector: pHO31; Site 1: Bst I; Constructed at The
               Institute for Genomic Research (TIGR), Rockville, MD.
               Genomic DNA isolated from broth cultures of E. histolytica
               using a method described by Clark and Diamond (Clark,
               C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
               method for isolate identification. Exp. Parasitol.
               77:450.). The DNA was mechanically sheared to give a
               tight size distribution (~2 kb). The v + i method used for
               the library construction is described in detail in Smith,
               H.O. and Venter, J.C. (Making small insert libraries for
               whole genome shotgun sequencing projects. In Genome
               Sequencing: A Practical Approach, eds. M. Vaudin and B.
               Borell, Oxford University Press, 1999)."
```

ORIGIN

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
35	119.00	839	62
61	41.47%	Conservative:	45
	24.03%	Mismatches:	103
	7.04%	Indels:	48
	8	Gaps:	13

US-10-724-972A-6352 (1-335) x AZ532836 (1-839)

QY 35 SerLysGluSerSerLysAspGlyValGluLeuLysHis-----GluGluGlyThrThr 52

Db 61 TCTAAGAGACAGAAAGCAATGATTGAATTCACAAATGCATTACAGACACATGAA 120

QY 53 LysValProLysHisProLysArgValVal----- 62

Db 121 AAGTCTTACCAATTTAAAAAGATTTATTGATCTTTAGACAATACTGTGAACCTGAA 180

QY 63 ----ValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAsp---ValLysProVal 80

Db 181 ACTGCTTTAAACCAAGTACAGTAAAGATTTATTACTGCTCCCTTGATGGAATTCGTGAACT 240

QY 81 GlylleAlaAspAsnLysLysAsnArgllelleLysProLeuHrgAspLyslle--- 99

Db 241 GGAATGGGAGATTTCATAAAGGAATTCGTTGTTAGAGAAATCTTTTGTGATTGTAT 300

QY 100 ---GlyLysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluLysSerLys 118

Db 301 AATGAAGACACACAAAGTATT-----AATCAAGGACACATTAATGCTTTTACTACTCT 351

QY 119 LeuLysProAspLeuIleAlaAspAsnArgHisLysGlyLysGlyLysAspLeu 138

Db 352 CTTAAAAAAGAA-----TACACTGTAGAAATTAAGAAAGTA 387

QY 139 AsnLysIle-----AlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGlu 156

Db 388 CAACAGTTATCAACAAGTACACAAAGAAAGAAAAAATTTGATCAAGAC----- 438

QY 157 AsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGlyLysLys 176

Db 439 -----TTATCTAATGCAGAAAAAAGCGCATCTACTTCTAATTTTAAAGAA 486

QY 177 ArgLeuGluGluHisAspLysLysIleGluGluTyrLysLysGluIleThrMetAspLys 196

Db 487 GCTATGGGAAGATAT---AAACGTATAGAAAAAGAAAGACAAAAATTTTAATGATCAT 543

QY 197 AsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsn 216

Db 544 TTACAAGTAACGTGTTATGCTTAACAGAGCTACTTTTGGT----- 582

QY 217 SerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspVal 236

Db 583 TCTATTATTGGGTCATTT-----GCTTTATTCTTTAAATCTACTGGTATGTTCAATGACG 636

QY 237 ThrLysGlyLeuSerLysTyrLeuLys---GlyProTyrLeuGlnMetAsnThrGluThr 255

Db 637 TCATCTGGA---GATAAGTTCAATTCATATGAGCCCTGGCTTAATAAATATCAAAAGTCT 693

QY 256 LeuSerGlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSerSer 273

Db 694 AGTCTAATATCCGAGAGAACTACGTGCTTTGATGACAAATATAAAACAAAGT 747

RESULT 13

LOCUS AZ538305/c

DEFINITION ENTPO68TF Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, genomic survey sequence.

ACCESSION AZ538305

VERSION AZ538305.1 GI:11143120

KEYWORDS GSS.

SOURCE Entamoeba histolytica

ORGANISM Entamoeba histolytica

REFERENCE 1 (bases 1 to 906)

AUTHORS Loftus,B., Van Aken,S. and Fraser,C.

TITLE Determination of clone end sequences from Entamoeba histolytica

JOURNAL HM1:IMSS sheared DNA library

COMMENT Unpublished (2000)

Contact: Brendan J Loftus

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0208

Fax: 301 838 3543

Email: bjoftus@igr.org

Clones are derived from the Entamoeba histolytica HM1:IMSS sheared DNA library

Seq primer: M13-Forward

Class: shotgun

High quality sequence start: 142

High quality sequence stop: 861.

Location/Qualifiers

1..906

/organism="Entamoeba histolytica"

/mol_type="genomic DNA"

/strain="HM1:IMSS"

/db_xref="taxon:5759"

/clone_lib="Entamoeba histolytica Sheared DNA"

FEATURES

source

/note="Vector: pHOS1; Site 1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of *E. histolytica* using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) *Entamoeba histolytica*: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Borell, Oxford University Press, 1999)."

ORIGIN

Alignment Scores:
Pred. No.: 0.0155 Length: 906
Score: 119.00 Matches: 74
Percent Similarity: 39.32% Conservative: 53
Best Local Similarity: 22.91% Mismatches: 129
Query Match: 7.04% Indels: 68
DB: 8 Gaps: 12

US-10-724-972A-6352 (1-335) x AZ538305 (1-906)

```
QY 22 AlaThrAlaAaCySGlyAsnAsnSerSerAsnSerSerLysGluSerSerLysAsp 41
DB 895 AGTACTTCTCAAAATAAATACTGACNAGACGAGTGGAAAGAAACAAACGCT-GAG 827
QY 42 GlyValGluIleLysHisGluGluGlyThrLysValProLysHisProLysArgVal 61
DB 826 GCAGAGAGAGAAAAAGGAA----- 806
QY 62 ValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGly 81
DB 805 -----ATGCAAGATAGTTTGAAGCA-----ATAAACAAAAAAGA 770
QY 82 IleAlaAspAsnLysAsnArgIleLysProLeuArgAspLysIleGlyLys 101
DB 769 CTTGAAGAAGTAAAAAAGAAAAATCAATACATCAGAA---AATCTAAAACGCTGCTAGC 713
QY 102 TyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLysPro 121
DB 712 CTTTCTTCC-----AAACAAAAAGAGCTTGAAGAAATAGAAAAACAAAAACAA 665
QY 122 AspLeuIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIle 141
DB 664 GAATTA-----GATGATCAAGCTAAAGAAAGATTGAAGAAAGAAAAACAAAGAAA 617
QY 142 AlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPhe 161
DB 616 GAAAAGAACTTAACCTGCACAGNAGAGAAAGAAAGAAAGAGAGAAATTTGAAAAAACT 557
QY 162 LysThrIleSerLysAlaLeuGlyLysGluGluGluGlyLys----- 175
DB 556 AAACAGGAAGAGAAAAAGAAAAAGAGATGAAGATTTAAACAAACAAAAAGAGAGAA 497
QY 176 -----LysArgLeuGluGluHisAspLysLysIle---GluGluTyrLysLysGlu 191
DB 496 GAAAATCTGAAAAGACTTGAGGAGATGAACGAAAAATAAAGAAAGAAAAAGAAAAAG 437
QY 192 IleThrMetAspLysAsnGlnLysValLeuProAlaValAlaAlaLysSerGlyLeuLeu 211
DB 436 ACTGAAGAAGAGAAAAAGAAAAAGATGTAAAAGAGTTTAAATCAAAACAA---GTAACA 380
QY 212 AlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAla 231
DB 379 ATTGAAGAGATACCTCTAATAAAGAAATAAACCCAGCAACAAAGAAAGAAACAAATT 320
QY 232 LeuSerAspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMet 251
DB 319 AAACAGCAAAATATTTCAAAAACAAAAAATGTGAGTATTAAACAAAGAGAAAAAGCAATA 260
```

```
QY 252 AsnThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsnLysAla 271
DB 259 GAAGAAGAA-----AGACAATAAGAGAAGAACAA-----ATAGAAGAGAAAAAG 212
QY 272 SerSerAsnGluProSerLysGluLeuGluLysAspProValTyrLysLysLeuAsn 291
DB 211 CAAACAAAGAAAAAGAAATTTATCACTTGAAGAAAAACAA----- 170
QY 292 AlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTyrPalaArgSerArgGly 311
DB 169 -----AAGAAGAACAGAGAGTTTGAACAAAGCAACTC----- 134
QY 312 LeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAspSerLys 331
DB 133 -----CAAGAAAAAACCACTTTAGAAAAAACTATTATTAGTAAG 98
QY 332 LysAspAsn 334
DB 97 AAAAGTAAT 89
RESULT 14
CL957361
LOCUS
DEFINITION
CL957361 1782 bp DNA linear GSS 21-SEP-2004
OsIFCC000458 Oryza sativa Express Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.
ACCESSION
CL957361
VERSION
CL957361.1 GI:52369736
KEYWORDS
GSS.
SOURCE
Oryza sativa (indica cultivar-group)
ORGANISM
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 1782)
AUTHORS
Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
Wong,G.K.S., Deng,X.W. and Wang,J.
TITLE
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
JOURNAL
Unpublished (2004)
COMMENT
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
FEATURES
Location/Qualifiers
source
1..1782
/organism="Oryza sativa (indica cultivar-group)"
/mo_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences "
```

ORIGIN

Alignment Scores:
Pred. No.: 0.0402 Length: 1782
Score: 119.00 Matches: 71
Percent Similarity: 40.27% Conservative: 76
Best Local Similarity: 19.45% Mismatches: 134
Query Match: 7.04% Indels: 84
DB: 9 Gaps: 14

US-10-724-972A-6352 (1-335) x CL957361 (1-1782)

```
QY 27 GlyAsnAsnSerSerSerAsnSerSerLysGluSerLysAspGlyValGluIleLys 46
DB 475 GGGAGCGATGGATCATCATCTCTTCTTCAGAAATCTGAGTTGACCAATGAGA 534
QY 47 HisGluGluGlyThrThrLysValProLysHisProLysArgValValLeuGluTyr 66
```



```
Db 702 ATTGAAAGACGAAAGAGGTTTAAAGAAAAACA-----AAAGATTGG 658
Qy 112 AenLeuGluGluIleSerLysLeuLysProAspLeuIleIleAlaAspAsnArgHis 131
Db 657 GAGGAAAAGACATTAAAGCTATAGAAACAGAAAT-----AAAAAC 616
Qy 132 LysGlyIleTyLysAspLeuAsn-----LysIleAlaProThrIleGluLeuLys 148
Db 615 AAAACCCCAATTAAGAAATGAATGATAGATGTAAATTCGAAGATAATAATACCAA 556
Qy 149 SerPheAspGlyAspTyAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeu 168
Db 555 -----GAAAGTATA 547
Qy 169 GlyLysGluGluGluGlyLysArgLeuGluGluHisAspLysLysIleGluGluTy 188
Db 546 ATGAAATATACAGAGAAACAAACAAATTAAGAGAA-----AAGATAGAGAGAA 496
Qy 189 -----LysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAla 206
Db 495 AAGATATAAGAGAAATCTGCTCTGAAATCA-----ATTACTGAA 451
Qy 207 LysSerGlyLeuLeuAlaHisProSerAsnSerTyValGlyGlnPheLeuSerGlnLeu 226
Db 450 TTACAAGGAAACTTAAAGAAATTAAGT-----ATGAAAGAAATG 412
Qy 227 GlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyLeuLysGly 246
Db 411 GAAGAAGACCAAGCGATTTCAGAGAAATGATGAAAAAGCAAAAGAAATAGTCAAAAA 352
Qy 247 ProTyLeuGlnMetAsnThrGluThrLeuSer-----GlnValAsnPro 261
Db 351 GAAATTGAAGAGAAATTCAGAAATGAAACACAAATAGAAAGATTCAAACGAATTAT 292
Qy 262 GluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLysLysGluLeu 281
Db 291 GAAAGAAATCAAGAAAGAAATGAAATTTAGAGAAAGAAATTTAAATTTACAGGAAAA 232
Qy 282 GlyLysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeu 301
Db 231 ATAATAGATAGAGAAAGAAATTCAGAGAAATTAATAAAGAGAGAAAAATC--- 175
Qy 302 AspArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGlu 321
Db 174 ---CGAGTATCAAGCTAATAAAGAGAGATGGAAGAAAGAAATGAGAAATGAGAGAA 118
Qy 322 LeuValGluLeuSerLysLysAspSerLys 331
Db 117 GAAATTGAAGGTTGAAAAAGAGAAACRAG 88

RESULT 16
AK029960
LOCUS AK029960 2740 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched
library, clone:4932411D20 product:similar to CTCL TUMOR ANTIGEN
SE2-2 (FRAGMENT) [Homo sapiens], full insert sequence.
ACCESSION AK029960
VERSION AK029960.1 GI:26325891
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
2
REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
```

```
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS
20499374 11042159
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,S., Nishine,T., Harada,A.,
Yanamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kaishiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913 11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2740)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan [E-mail:genome-res@gsc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216]
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/.
Location/Qualifiers
1..2740
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:4932411D20"
/db_xref="caxon:10090"
/clone="4932411D20"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
39..2465
/note="unnamed protein product; putative
similar to CTCL TUMOR ANTIGEN SE2-2 (FRAGMENT) [Homo
sapiens] (SPTR|Q9H2G6, evidence: FASTY, 83.6%ID,
```

ORIGIN

```
Alignment Scores:
Pred. No.: 0.0265 Length: 1005
Score: 117.50 Matches: 61
Percent Similarity: 37.9% Conservative: 59
Best Local Similarity: 19.30% Mismatches: 125
Query Match: 6.95% Indels: 71
DB: 5 Gaps: 11

US-10-724-972A-6352 (1-335) x BQ065769 (1-1005)

QY 27 GlyAsnAsnSerSerSerAsnSerSerLysGluSerSerLysAspGlyValGluLeuLys 46
DB 15 GGAGAGAAACCAAGCCCTGTCCTGGCTCGGGCCCTTGAAGAGCCCTTGAAGCCCAA 74

QY 47 HisGlu---GluGlyThrThrLysValProLysHisProLysArgValValValLeuGlu 65
DB 75 GAGGAACCTCGAGCGGACCAACAAA-----ATGCTCAAA 107

QY 66 TyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyLeAlaAspAsp 85
DB 108 GCCGAATGGAAGACCTGGTCAGCTCCAAGGATGAGTGGGCAAGACGTCATGAGCTG 167

QY 86 AsnLysLysAsnArgIleLysProLeuArgAspLysIleGlyLysTyrThrSerVal 105
DB 168 GAGAAGTCCAAGCGGCCCTGGAG----- 191

QY 106 GlyThrArgLysGlnProAsnLeuGluGluLysSerLysLeuLysProAspLeuLys 125
DB 192 ---ACCCAGATGGAGAGATGAAGACGCGCTGGAAGAGCTGGAGGACGAGCTGCAAGCC 248

QY 126 AlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIle 145
DB 249 ACGGAGGAC-----GCCAAACTCGCGGTGGAAGTC 278

QY 146 GluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSer 165
DB 279 AACATGCGAGGCGCTCAAGCGCCAGTTCGAAGGATCTCCAAAGCC----- 323

QY 166 LysAlaLeuGlyLysGluGluGluGlyLysLysArgLeuGluGluHisAspLysLysIle 185
DB 324 ---CGGAGCAGCAGATGAGGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 371

QY 186 GluGluTyrLysLysGluLeuThrMetAspLysAsnGlnLysValLeuProAlaValAla 205
DB 372 CACGAGTATGAGACGGAAGTGAAGACGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 431

QY 206 AlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGln 225
DB 432 AAGAAG-----AAGCTGGAAGGGGAC 452

QY 226 LeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSer-----Lys 242
DB 453 CTGAAGACCTGGAGCTTCAGCGGACCTTCGCATCAAGGGGAGGAGGAGGAGGAGGAGGAG 512

QY 243 TyrLeuLysGlyProTyrLeuGlnMetAsn-----ThrGluThrLeuSerGlnValAsn 260
DB 513 CAGCTACGCAAACTCGAGCTCAGATGAAGACTTTCGAAGAGAGCTGGAAGATCCCGT 572

QY 261 ProGluArgMetPheIleMetThrAsnLysAlaSerAsnGlnProSerLeuLysGlu 280
DB 573 GCCTCCAGAGATGAGATCTTTGCCACA---GCCAAAGAGATGAGAAGAAAGCCAAAGAGC 629

QY 281 LeuGluLysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIle 300
DB 630 TTGAAGACGAGACTTATGTCAGCTCAAGAGGACCTCGCGCGCTGAGAGG----- 680

QY 301 LeuAspArgAspLeuTrpAlaArgSerArgGlyLeuLysSerSerGluGluMetAlaLys 320
DB 681 -----GCTCGCAACAAAGCGGACCTCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 722

QY 321 GluLeuVal---GluLeuSerLysLysAspSerLysLysAspAsnLys 335
DB 723 GAGCTGGCCAGTAGCGTGTGGGAAAGGAGCGCACTCCAGGACGAGAAG 770
```

```
RESULT 18
CF109973
LOCUS 1065 bp mRNA linear EST 23-JUL-2003
DEFINITION Shultzomica03224 Rat lung airway and parenchyma cDNA libraries
Rattus norvegicus cDNA clone Contig2841 5', mRNA sequence.
ACCESSION CF109973
VERSION CF109973.1 GI:33165360
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1065)
AUTHORS Shultz,M.A., Zhang,L., Gu,Y.-Z., Baker,G.L., Fannuchi,M.V.,
Padua,A.M., Gurske,W.A., Morin,D., Penn,S.G., Jovanovich,S.B.,
Plopper,C.G. and Buckpitt,A.R.
TITLE Gene expression analysis in response to lung toxicants: I.
JOURNAL Sequencing and microarray development
COMMENT Am. J. Respir. Cell Mol. Biol. 30 (3), 296-310 (2004)
Contact: Shultz MA
Dept. of Molecular Biosciences, School of Veterinary Medicine
University of California, Davis
1311 Haring Hall, One Shields Avenue, Davis, CA 95616, USA
Tel: 530 752 0793
Fax: 530 752 4698
Email: mashultz@ucdavis.edu
Average Phred score is 20 or better. All poor quality data (Phred <
20) and vector/linker sequence has been removed.
High quality sequence stop: 1065.
FEATURES
Location/Qualifiers
source
1..1065
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="Contig2841"
/sex="male"
/tissue_type="airway or parenchyma"
/dev_stages="adult"
/clone_lib="Rat lung airway and parenchyma cDNA libraries"
/note="Organ: lung; Vector: pGEM-11zf(-); Site.1: Eco RI;
Site 2: Not I; mRNA was isolated from microdissected rat
lung airways and parenchyma tissues."
ORIGIN
Alignment Scores:
Pred. No.: 0.0327 Length: 1065
Score: 117.00 Matches: 51
Percent Similarity: 37.56% Conservative: 29
Best Local Similarity: 23.94% Mismatches: 59
Query Match: 6.92% Indels: 74
DB: 7 Gaps: 9
US-10-724-972A-6352 (1-335) x CF109973 (1-1065)
QY 28 AsnAsnSerSerSerAsnSerSerLysGluSerSerLys----- 40
DB 413 AACAAATGAAGAAGATCTGCCTCAAAACAAGATAAAGAGGAGGAGGAGGAGGAGGAGGAG 472
QY 41 -----AspGlyValGluLysHisGluGlyThrThrLys 53
DB 473 AAAACAAGTTTGTAGATGAGATAGTAGAAGATTTGGAGATTAAGATCAAAATCAAAA 532
QY 54 ValProLysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeuVal 73
DB 533 AAGCTGCCAGACCTTAAGTCAAGAGTGTCTCTCT----- 568
QY 74 AlaLeuAspValLysProValGlyLeAlaAspAsnLysLysAsnArgIleLys 93
DB 569 -----GGGAGTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 601
```

QY 94 ProLeuArgAspLysIleGlyThrSerValGlyThrArgLysGlnProAsnLeu 113
Db 602 -----TCTAAAGAGTAAAGAGCCAGCAATCAACTAAAGAGCGGATGGTCG 652
QY 114 GluGluIleSerLysLeuLysProAspLeuIleAlaAspAsnAsnArg----- 130
Db 653 GAAGAGGATCAA-----GACATAGCAAGAAAGTAA 685
QY 130 ----- 130
Db 686 GAACGTTTCGAGAGTAAACTCTCAGGTGAAGTGGTGGTCTGATCAATTTTTCAG 745
QY 131 ---HisLysGlyIleThrLysAsp---LeuAsnLysIleAlaProThrIleGluLeuLys 148
Db 746 TCCAGAAAGGAGCAGAAAGAAAGTCAAGAACAGTCACTTCCATCAGTACAGTGG 805
QY 149 SerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSer-----Lys 166
Db 806 AATGAGGACGATGACTCTTCTTCAAAATT-----AAGACGTTGGCCAGAGAG 856
QY 167 AlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAspLysLysIleGlu 186
Db 857 GCAGAAAGAGAGCGTGCAGAGAGAAAGAGAG---GAAGAGGAGAAAGCAAAGTTGCGA 913
QY 187 GluTyrLysLysGluIleThrMetAspLysAsnGlnLys 199
Db 914 AAGGTGAAGAGAAAGAGAGCTAGAGAGGCGAGGAG 952

RESULT 19
LOCUS AY407955
DEFINITION Homo sapiens MYH11 gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY407955
VERSION AY407955.1 GI:39763926
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.

TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
1..5504
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>5504
/gene="MYH11"
/locus_tag="HCM3063"

ORIGIN
Alignment Scores:
Pred. No.: 0.328 Length: 5504
Score: 117.00 Matches: 73
Percent Similarity: 35.32% Conservative: 63

Best Local Similarity: 18.96% Mismatches: 146
Query Match: 6.92% Indels: 103
DB: 9 Gaps: 11
US-10-724-972A-6352 (1-335) x AY407955 (1-5504)
QY 22 AlaThrAlaAlaCysGlyAsnAsnSerSerSerLysGluSerSerLysAsp 41
Db 4137 GCTGCAGGACTTTTGCAGCACCGTGGAAAGCTCTGGAAGAGGAGAGAGGTTCCAGAA 4196
QY 42 Gly-ValGlu-----IleLysHisGluGlyThrThrLysValProLysHisPr 58
Db 4197 GGAGATCGAAGAACCTCACCAGCAGTACGAGAGAGAGGCGCGCTTATGATAAAGTGA 4256
QY 58 oLysArgValValValLeuLysThrSerPheValAspAlaLeuAlaLeuAspVal 78
Db 4257 AAGACCAAGAACAGGCTTCAGCAGGAGCTGGACCACTGGTGTGTTGATTTCGACAACCA 4316
QY 78 sProValGlyIleAlaAspAsnLysLysAsnArgIleLysProLys----- 95
Db 4317 GCGCAACTCGTGTCCAACTCGAAAGAGAGAGGAAATTTGATCAGTTCTTACGCCA 4376
QY 96 -----Ar 96
Db 4377 GGAGAAACATCTCTTCAATACCGGATGAGAGGAGCAGAGCTGAGGAGAGAGCCAG 4436
QY 96 sAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluL 116
Db 4437 GGAGAGGAAACCAAGGCCCTGTCTCGGCGCTTGAAGAGGCGCTTCGAAGCCAA 4496
QY 116 sSerLysLysLys-----ProAspLeuIleAla 126
Db 4497 AGAGAACTCGAGCGGAGCAACAAATGCTCAAAGCCGAAATGGAAGACCTGTCAGCTC 4556
QY 126 sAspAsnAsnArgHisLysGlyIleTyrLys----- 136
Db 4557 CAAGATGACGTGGCAAGAGCTCCTAGCTGAGAGAGTCCAAGCGGCGCTCGAGAC 4616
QY 137 -----Ab 137
Db 4617 CCAGATGAGGAGATGAAGACGACGCTGGAAGAGCTGGAGGAGGAGTGCAGACCCAG 4676
QY 137 pLysAsnLysIleAlaProThrIleGluLysSerPheAspGlyAspTyrAsnGlu 157
Db 4677 GGACCCAACTGCGCGCTGGAAGTCAACATGCGCGCTCAAGGCGCAGTTCGAAGGGA 4736
QY 157 nIleAspAlaPheLysThrIleSerLysAlaLeuLysGluGluGlyLysAsr 177
Db 4737 TCTCAAGCC-----CGGACGAGCAGAGATGAGAGAA 4769
QY 177 gLeuGluHisAspLysLysIleGluGluTyrLysLysGluIleThrMetAspLys 197
Db 4770 GAGGAGGCACTGCAGACAGCTTCACAGATATGAGACGGAACCTGGAAGAGGAGGAA 4829
QY 197 nGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsn 217
Db 4830 GCAACGTCCTCGCAGCTGCAGCAAGAG----- 4860
QY 217 rTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspVal 237
Db 4861 -----AAGCTGAAGGGGACCTGAAAGACCTGAGGCTTCAGGCGGCTCTGCCAT 4910
QY 237 rLysGlyLeuSer-----LysTyrLysGlyProTyrLeuGlnMetAsn----- 252
Db 4911 CAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4970
QY 253 -ThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsnLysAla 272
Db 4971 TCAAGAGAGAGCTGGAAGATGCGCGCTCCAGAGATGAGATCTTTGCCACA---GCCAA 5027
QY 272 rSerAsnGluProSerLeuLysGluLeuGluLysAspProValTrpLysLysLeuAsn 292
Db 5028 AGAGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5087


```

/organism="Cossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Eb013122"
/tissue_type="floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_lib="GR_Eb"
/notes="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; Library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."

ORIGIN
Alignment Scores:
Pred. No.: 0.0292 Length: 816
Score: 116.00 Matches: 68
Percent Similarity: 37.89% Conservative: 40
Best Local Similarity: 23.86% Mismatches: 119
Query Match: 6.86% Indels: 58
DB: 7 Gaps: 11

US-10-724-972A-6352 (1-335) x CO113362 (1-816)
QY 84 AspAspAsnLysGlnProLeuArgAspLysValleGlyLysTyrThr 103
DB 10 GAAGAGGGAAGAAGAA-----AAGCCG---AAGACAAAGGCAAGGAGAAG 57
QY 104 SerValGlyThrArgLysGlnProAsnLeuGluLeuSerLysLeu----- 119
DB 58 AAGAAGAAAGAAAGAAAGGATGAGTGAAGAGCTGAAGATTTAGACGGGGAAG 117
QY 120 -----LysProAspLeuIleAlaAspAsnAsnArgHisGlyGlyLeuTyrLys 136
DB 118 AAGAAAAAGAAAGAAAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 168
QY 137 AspLeuAsnLysIleAlaProThrIleGluLeu----- 147
DB 169 AAGAAGGATGAGGTAGAGAGAGTGAAGAAATTAGAGAGGAGGAGGAGGAGGAG 228
QY 148 -----LysSerPheAspGlyAspTyrAsn 155
DB 229 GACAAAGAGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 288
QY 156 GluAsnIleAspAlaPheLysThrIleSerLys-----AlaLeuGlyLysGlu 171
DB 289 GAAGAAATTAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 348
QY 172 GluGluGlyLysLysArgLeuGlu---GluHisAspLysIleGluGluTyrLysLys 190
DB 349 AAGAAAAAGAAAGAAAGGATGAGGTAGATGAAGATGAAGATTTAGAGAGGAG 408
QY 191 GluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeu 210
DB 409 GAAGAAGAAAGAAAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 468
QY 211 LeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGlu 230
DB 469 GTAGAAGAGCTCAAGAA-----TTAGAAGAGGGAATTAAGAA 507
QY 231 AlaLeuSerAspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGln 250
DB 508 AAGAAGAAAGAAAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 567
QY 251 MetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsnLys 270
DB 568 GAAGAAGCTGAAGATTTAGAGAGGGAATAAAGAAAGAAAGAAAGAAAGAGG 627
QY 271 AlaSerSerAsnGluProSerLeuLysGluLeuGluLysAspProValTyrLysLysLeu 290
DB 628 GAGAAGAGAAAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 678
QY 291 AsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArg 310

```

```

Db 679 GAAGAGGGAAGAAAGAAAG-----AAGAGAGACAAAGAGCGCAAGGAGAAG 726
QY 311 GlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAspSer 330
DB 727 -----AAGAAAGAAATGAAGAGGATGAGGTAGAGGTGAAGAGAAAGAGAG 774
QY 331 LysLysAspAsnLys 335
DB 775 AAGAAGGAGAGAGAG 789

RESULT 24
LOCUS A2549761
DEFINITION ENTFFK51TF Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, genomic survey sequence.
ACCESSION A2549761
VERSION A2549761.1 GI:11174477
KEYWORDS GSS.
SOURCE Entamoeba histolytica
ORGANISM Entamoeba histolytica
REFERENCE 1 (bases 1 to 889)
AUTHORS Loftus B., Van Aken, S. and Fraser, C.
TITLE Determination of clone end sequences from Entamoeba histolytica HMI:IMSS sheared DNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
7712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjloftus@tigr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 7
High quality sequence stop: 852.

FEATURES
source
1..889
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/notes="Vector: pHO51; Site 1: Bet I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999)."
```

```

ORIGIN
Alignment Scores:
Pred. No.: 0.0329 Length: 889
Score: 116.00 Matches: 63
Percent Similarity: 42.75% Conservative: 49
Best Local Similarity: 24.05% Mismatches: 75
Query Match: 6.86% Indels: 76
DB: 8 Gaps: 14

US-10-724-972A-6352 (1-335) x A2549761 (1-889)
QY 86 AsnLysLysAsnArgIleIleLysProLeuArgAsp-----LysIleGlyLysTyrThr 103

```

```
Db 191 AATAAGAGAAATAAAGAGATTAAATGATGGAAGAAAGAAATTAATAAATTCAAAAGAGAAAC 250
Qy 104 SerValGlyThrArgLysGlnProAsnLeuGluGluLeuSerLysLeuLysProAspLeu 123
Db 251 ACTCTTCCAAATTCACCAACCAAGATGATGAAAGAAAGAAATGAAA----- 301
Qy 124 IleIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaPro 143
Db 302 -----GAAATCGTAGA-----ACTATAGTCCA 325
Qy 144 ThrIle-----GlueLysSerPheAspGlyAspTyrAsnGluAsnIle 158
Db 326 AGGATGATGGGTTAGGAACAGAAATTAACAA-----AATAAGAAAGTAGAAGAAAAAAT 382
Qy 159 AspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGly----- 174
Db 383 AGAAGATAGAGAAATTAACAAGAAATTTAGAGTTAGAACACAAAGAAACAAATCAACTT 442
Qy 175 LysLysArgLeuGluGluHisAspLysLysIleGluGluTyrLysLysGluIleThrMet 194
Db 443 AAAGAAAGTTAGAGACAGCAAGAAATCAATTTGAAGAAATGAAGAAATTAATAAAA 502
Qy 195 AspLys-----AsnGlnLysValLeuProAlaValAlaLysSer 208
Db 503 GAAAGGAAGAGAGTTTGAAGAGAAATTAATGAGAG----- 535
Qy 209 GlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer-----GlnLeu 226
Db 536 -----AATAATATACAATTAATGAATGAATCAATTTTGAATTA 577
Qy 227 GlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGly 246
Db 578 GAAAGAAAGAAAA---GATGAAGAAATTAATAA----- 610
Qy 247 ProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheIle 266
Db 611 -----CTAAAAAGTTCAATTGAACAA----- 631
Qy 267 MetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLysAspProVal 286
Db 632 CAACATATCAAAATAGAACAACTCAGTTAGAAATTAATAAAGCTTTGAGGAATTAATAA 691
Qy 287 ---TPlLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeu 305
Db 692 GAAAGTGAAGAAACAAATGAATCAAAACAAAGAAATAGAAAGATTAAATAAGA-AAT 750
Qy 306 TrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeu 325
Db 751 TGG-----AGTTTAAAGACACTGAACATGAAGAAAGAGTAAAGAAATGAACCTTCTT 804
Qy 326 SerLys 327
Db 805 TTGAAA 810

RESULT 25
CR650378
LOCUS CR650378 1483 bp mRNA linear HTC 18-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR650378
VERSION CR650378.1 GI:51146823
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
1 (bases 1 to 1483)
REFERENCE
Genoscope.
AUTHORS Direct Submission
TITLE Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
JOURNAL i 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
```

```
(E-mail : seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
Location/Qualifiers
source 1. 1483
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/tissue_type="Liver"

ORIGIN
Alignment Scores:
Pred. No.: 0.0675 Length: 1483
Score: 116.00 Matches: 62
Percent Similarity: 41.48% Conservative: 50
Best Local Similarity: 22.96% Mismatches: 88
Query Match: 6.86% Indels: 70
DB: 3 Gaps: 14

US-10-724-972A-6352 (1-335) x CR650378 (1-1483)
Qy 50 GlyThrLysValProLysHisProLysArgValValLeuGluTyrSerPheVal 69
Db 499 GGTTCACCTCGATCCCAAAATCCAGAA----- 528
Qy 70 AspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsnLysLysAsn 89
Db 529 -----CTGCTCTGGAGCTCACTCTCTCTCGGTATTTGAGACG---GCCGGA 576
Qy 90 ArgIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLys 109
Db 577 GGCCTCATGACGGTCTG-----ATCAAGCGCAACACCATCTCTCCACGACG 627
Qy 110 GlnProAsnLeuGluLysSerLysLysProAspLeuIleAlaAspAsn 129
Db 628 ACCCAGACCTTCACACCTACTCTGACAAACGACCGCGTCTCATCCAGGTTTACGAG 687
Qy 130 ArgHisLysGlyIleTyrLysAspLeuAsnLysIle----- 141
Db 688 GGTGAGCGTGCATGACGAGGACAAATCTGCTGGCAAGTTCGAGTTGACGGGCATC 747
Qy 142 -----AlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsn 155
Db 748 CCCCTCTCCCTCTGCGGTCTCCAGATCAGGTG---ACTTTTGACATCGACGCCAAC 804
Qy 156 -----GluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 173
Db 805 GGCATCATGAACGTGTCTGCC-----GTGACAAAGCACTGGCAAGAAACAAAG 855
Qy 174 -----GlyLysLysArgLeuGluGluHisAsp-----Lys 183
Db 856 ATCACCATTACCAATGACAAAGGTCTCTAGTAGGAGGACATTGACCGCATGTCGAG 915
Qy 184 LysIleGluGluTyrLysLysLysLysIleThrMetAspLysAsnGlnLysValLeuProAla 203
Db 916 GAAGCTGAGAGTACAGGCCGCAAGACGCTCCGCGGTGACAAAG----- 960
Qy 204 ValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeu 223
Db 961 GTGTGCGCAAGAAATTTGAGTCTGCTGCTTCAAC----- 999
Qy 224 SerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGly---LeuSerLys 242
Db 1000 -----ATGAGTCAACCTGGAGACGAGAACTCGCTGGGAAGATCATGTGAC 1047
Qy 243 TyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGlu 262
Db 1048 GATGACAGCAGAAAGATTTTGGACAAAGTGCAACGAGGTTCATCAGCTGCTCGACAGAAT 1107
Qy 263 ArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGlu 282
Db 1108 CAG-----ACTGCCGAGAAAGACGAGTACGAGCATCAGACGAGGAGGCTGGAG 1155
```

```

QY      283 Lys-----AspProValTrpLysLysLeu 290
      |||
      ::|||:::
Db      1156 AAGTGTGCAACCCCATCATCAAGATG 1185

RESULT 26
AL854856/c
LOCUS   AL854856 XGC-egg Xenopus tropicalis cDNA clone Tegg004f24 3', mRNA
DEFINITION
ACCESSION
VERSION  AL854856
KEYWORDS EST.
SOURCE   Xenopus tropicalis (western clawed frog)
ORGANISM
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
          Xenopodinae; Xenopus; Silurana.
REFERENCE
  1 (bases 1 to 673)
AUTHORS  Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
TITLE    Sanger Xenopus tropicalis EST project 2001 (11_2003)
JOURNAL  Unpublished (2003)
COMMENT  On Sep 15, 2002 this sequence version replaced gi:22875075.
          Contact: Taylor R
          Sanger Institute
          Hinxton, Cambridgeshire, CB10 1SA, UK
          Email: trop@sanger.ac.uk
          Sanger Xenopus tropicalis EST project 2001
          TROPICALIS SEQUENCE ID: Tegg004f24.q1kT7
          Sequencing primer: 77
          This sequence is from a Xenopus Gene Collection (XGC) library
          constructed by Aaron M. Zorn.
          cDNA was oligo dT primed from 5ug of poly A+ RNA from egg.
          EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
          5' end and NotI at the 3' end.
          Vector: pCS107; Site_1: EcoRI; Site_2: NotI
          Host: Escherichia coli XL1-blue.
          Location/Qualifiers
            1..673
              /organism="Xenopus tropicalis"
              /mol_type="mRNA"
              /db_xref="taxon:8364"
              /clone="Tegg004f24"
              /dev_stage="egg"
              /lab_host="Escherichia coli XL1-blue"
              /clone_lib="XGC-egg"
              /notes="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
              was oligo dT primed from 5ug of poly A+ RNA from egg.
              EcoRI-NotI cut cDNA was then ligated into pCS107 with
              EcoRI at the 5' end and NotI at the 3' end"

ORIGIN
Alignment Scores:
Pred. No.: 0.0254 Length: 673
Score: 115.50 Matches: 60
Percent Similarity: 39.38% Conservative: 42
Best Local Similarity: 23.17% Mismatches: 106
Query Match: 6.83% Indels: 51
DB: 1 Gaps: 8

US-10-724-972a-6352 (1-335) x AL854856 (1-673)

QY      82 IleAlaAspAsnLysLysAsnArgIleLeLysProLeuArgAspLysIleGlyLys 101
      |||
      ::|||:::
Db      643 ATACATGACCTGAGAGCGATCAGTCATGTGAGAGCCCAAGAGAGATACAGGAA 584

QY      102 TyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLysPro 121
      |||
      ::|||:::
Db      583 GATACAAAAGAAATTAGTGAAGAAAGCAATACGCTCTTGGAAACATGAAGGAAAG--- 527

QY      122 AspLeuIleIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIle 141
      |||
      ::|||:::
Db      526 -----AACCAAGGACTTAAAGATGTAGAGAGAGCAACTTAAACAAATA 485

```

```

QY      142 AlaProThrIleGlu-----LeuLysSerPheAspGlyAspTyrAsnGluAsnIleAsp 159
      |||
      ::|||:::
Db      484 ACAAGTTTCATTGAAGAAACAGAGAGAGATTTTCTCAGTTGGACTTACAAGATGTGAC 425

QY      160 AlaPheLysThrIleSerLysAlaLeuGlyLysGluGluGluGluLysArgLeuGlu 179
      |||
      ::|||:::
Db      424 ACTGTGAGAACTGAACACTCAAAAGCAAGTTTAAAAAAGTTCGCAAAACAGCTTCAA 365

QY      180 GluHisAspLysIleGluGlyTyrLysLysGluIleThrMetAspLysAsnGlnLys 199
      |||
      ::|||:::
Db      364 AAAGACAAAGAAAGGTGGAAGAAATAAAGAT-----GTGCTGCTAATAGCCAAAG 311

QY      200 ValLeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrVal 219
      |||
      ::|||:::
Db      310 ATTATTACTGAGAAACAAATAAAGAGGATCTGCTGGAA----- 272

QY      220 GlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAspValThrLysGly 239
      |||
      ::|||:::
Db      271 -----AAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 224

QY      240 LeuSerLysTyrLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnVal 259
      |||
      ::|||:::
Db      223 CTTAAGAGGAAACGCAAGGT-----TTGCAA----- 197

QY      260 AsnProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLys 279
      |||
      ::|||:::
Db      196 -----GAGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 179

QY      280 GluLeuGluLysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAsp 299
      |||
      ::|||:::
Db      178 GAAAGAGAACTAATGGAATAAGCAAACTGTAATAGAGGCTGCTTCA---AAATGGAT 122

QY      300 IleLeuAspArgAspLeu-----TrpAlaArgSerArgGlyLeuIleSerSerGlu 316
      |||
      ::|||:::
Db      121 GTTGCTCAGTCAGAGCTGATATATATTTGAGCCGCCATAACTGCGTGTGCCAATT 62

QY      317 GluMetalLysGluLeuValGluLeuSerLysLysAspSerLysLysAspAsnLys 335
      |||
      ::|||:::
Db      61 AATAAGGCAAGAGAGCTTTAAACACAGCTTCAGCAACTCTAAAGAGAGAGAGAGAGAG 5

RESULT 27
BW253010 684 bp mRNA linear EST 09-NOV-2002
LOCUS   BW253010 Nori Satoh unpublished cDNA library, tailbud embryo Ciona
DEFINITION
          intestinalis cDNA clone citb094p15 5', mRNA sequence.
ACCESSION  BW253010
VERSION    BW253010
KEYWORDS   EST.
SOURCE     Ciona intestinalis
           Ciona intestinalis
           Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
           Phlebobranchia; Cionidae; Ciona.
REFERENCE  1 (bases 1 to 684)
AUTHORS   Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
TITLE     Expressed genes in Ciona intestinalis (2002c)
JOURNAL   Unpublished (2002)
COMMENT   Contact: Nori Satoh
           Department of Zoology
           Kyoto University
           Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
           Tel: 81-75-753-4081
           Fax: 81-75-705-1113
           Email: satoh@scidian.zool.kyoto-u.ac.jp.
           Location/Qualifiers
             1..684
               /organism="Ciona intestinalis"
               /mol_type="mRNA"
               /db_xref="taxon:7719"
               /clone="citb094p15"
               /tissue_type="whole animal"
               /dev_stage="tailbud embryo"
               /clone_lib="Nori Satoh unpublished cDNA library, tailbud
source

```

```

ORIGIN
embryo"

Alignment Scores:
Pred. No.: 0.0259 Length: 684
Score: 115.50 Matches: 65
Percent Similarity: 41.04% Conservativeness: 38
Best Local Similarity: 25.90% Mismatches: 91
Query Match: 6.83% Indels: 57
DB: 14 Gaps: 14

US-10-724-972A-6352 (1-335) x BW253010 (1-684)
Qy 69 ValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsnLysLys 88
Db 8 GTTCAGGATCTCTCTGTGGATGTGCACCTTTATCTCTGTGATTGAGACC---GCT 64

Qy 89 AsnArgIleLeuProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArg 108
Db 65 GGAGGAGTGTGATGACAGCTTTG-----ATCAAGAGAGAAACACGACCGTCCCAACCAAA 115

Qy 109 LysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsn 128
Db 116 CAATCCCAAAATATTCACCACTTACTCGGACAAACCAACGAGGTGTGTGATCCAAAGTGATC 175

Qy 129 AsnArgHisLeuGlyIleTyrLysAspLeuAsnLysIle----- 141
Db 176 GAGGTGTAACGTACGATGACCAAGAGACAAACATATCTTGTAAAGTTCGAGCTCTCCGGC 235

Qy 142 -----AlaProThrIleGluLeuLysSerPheAspGlyAspTyr 154
Db 236 ATTCCCTCCGACCTCGTGGAGTTCACAGATTGAGGTG---ACGTTGATATCATGCT 292

Qy 155 AsnGlu-----AsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGlu 172
Db 293 AACGCTATCTGAATGTATCTGCT-----GTTGACAAAGCAGCACTGGAAGGAGAA 343

Qy 173 Glu-----GlyLysLysArgLeuGluGluHisAsp----- 182
Db 344 AGATCATCAATCAACCAACGACAAAGGTGCTCTTAGCAAGAGATATTGACGGGATGGT 403

Qy 183 LysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuPro 202
Db 404 AACGAAGCGGAGAAATATAAGAGAA-----GATGAGAAGCAGAAAG-----GAG 448

Qy 203 AlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe 222
Db 449 AAGATTCAAGCGAAGAACCGACTCGAATCTTATGCGTTCAAC----- 490

Qy 223 LeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGly---LeuSer 241
Db 491 -----CTTAATCAACAGTGGAGGATGACAAAGTGAAGATAAAATTTC 535

Qy 242 LysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnPro 261
Db 536 GACGAAGATAAATCAGCGATTCTTAACAAAGCTAAAGAAAGTTTGGACTGGTGGAAAT 595

Qy 262 GluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu 281
Db 596 AACCAAG-----ACTCCCAAGAAAGGAGCAATATGAATTCACAAAAAGAACTT 643

Qy 282 GluLys-----AspProValTrpLysLysLeu 290
Db 644 GAGAAATTTGCCAACCTTATTATGACAAAACCTT 676

RESULT 28
LOCUS BW254793
DEFINITION BW254793 Nori Satoh unpublished cDNA library, tailbud embryo Ciona
intestinalis cDNA clone citb073p02 5', mRNA sequence.
ACCESSION BW254793
VERSION BW254793.1 GI:24834711
KEYWORDS EST.

```

```

SOURCE ORGANISM
Ciona intestinalis
Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cionidae; Ciona.

REFERENCE
1 (bases 1 to 690)
Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
Expressed genes in Ciona intestinalis (2002c)
Unpublished (2002)
Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoheascidian.zool.kyoto-u.ac.jp.

FEATURES
Location/Qualifiers
1..690
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="citb073p02"
/tissue_type="whole animal"
/dev_stage="tailbud embryo"
/clone_lib="Nori Satoh unpublished cDNA library, tailbud
embryo"

ORIGIN

Alignment Scores:
Pred. No.: 0.0263 Length: 690
Score: 115.50 Matches: 65
Percent Similarity: 41.04% Conservativeness: 38
Best Local Similarity: 25.90% Mismatches: 91
Query Match: 6.83% Indels: 57
DB: 14 Gaps: 14

US-10-724-972A-6352 (1-335) x BW254793 (1-690)
Qy 69 ValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsnLysLys 88
Db 19 GTTCAGGATCTCTCTGTGGATGTGCACCTTTATCTCTGTGATTGAGACC---GCT 75

Qy 89 AsnArgIleLeuProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArg 108
Db 76 GGAGGAGTGTGATGACAGCTTTG-----ATCAAGAGAGAAACACGACCGTCCCAACCAAA 126

Qy 109 LysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsn 128
Db 127 CAATCCCAAAATATTCACCACTTACTCGGACAAACCAACGAGGTGTGTGATCCAAAGTGATC 186

Qy 129 AsnArgHisLeuGlyIleTyrLysAspLeuAsnLysIle----- 141
Db 187 GAGGTGTAACGTACGATGACCAAGAGACAAACATATCTTGTAAAGTTCGAGCTCTCCGGC 246

Qy 142 -----AlaProThrIleGluLeuLysSerPheAspGlyAspTyr 154
Db 247 ATCCCTCCGACCTCGTGGAGTTCACAGATTGAGGTG---ACGTTGATATCATGCT 303

Qy 155 AsnGlu-----AsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGlu 172
Db 304 AACGCTATCTGAATGTATCTGCT-----GTTGACAAAGCAGCACTGGAAGGAGAA 354

Qy 173 Glu-----GlyLysLysArgLeuGluGluHisAsp----- 182
Db 355 AGATCATCAATCAACCAACGACAAAGGTGCTTTAGCAAGAGATATTGACGGGATGGT 414

Qy 183 LysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuPro 202
Db 415 AACGAGGCGGAGAAATATAAGAGAA-----GATGAGAAGCAGAAAG-----GAG 459

Qy 203 AlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe 222
Db 460 AAGATTCAAGCGAAGAACCGACTCGAATCTTATGCGTTCAAC----- 501

```

```

QY 223 LeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGly---LeuSer 241
Db 502 -----CTTAAATCAACGGTGGAGATGACAAAGTGAAGATGAAGATTCA 546
QY 242 LysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnPro 261
Db 547 GACGAAGATAAGTCAGCGATTCTTAACAAGCTAAAGAAGTTTGGACTGGTTGGAAT 606
QY 262 GluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu 281
Db 607 AACACAG-----ACTCGGAAAGAGGACGAATATGAATTCACAAAAAGAACTT 654
QY 282 GluLys-----AspProValTrpLysLysLeu 290
Db 655 GAGAAAAATTGCCAACCCCTATTATGACTAAACTT 687

RESULT 29
BW222839
LOCUS BW222839 692 bp mRNA linear EST 06-NOV-2002
DEFINITION BW222839 Nori Satoh unpublished cDNA library, larva Ciona
intestinalis cDNA clone cilv051j23 5', mRNA sequence.
ACCESSION BW222839
VERSION BW222839.1 GI:24742218
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cionidae; Ciona.
REFERENCE 1 (bases 1 to 692)
AUTHORS Satou,Y., Shin-i,T., Kohara,Y. and Satoh,N.
TITLE Expressed genes in Ciona intestinalis (2002c)
JOURNAL Unpublished (2002)
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoheascidian.zool.kyoto-u.ac.jp.

FEATURES
source
1..692
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="cilv051j23"
/tissue_type="whole animal"
/dev_stage="larva"
/clone_lib="Nori Satoh unpublished cDNA library, larva"

ORIGIN
Alignment Scores:
Pred. No.: 0.0264 Length: 692
Score: 115.50 Matches: 65
Percent Similarity: 41.04% Conservative: 38
Best Local Similarity: 25.90% Mismatches: 91
Query Match: 6.83% Indels: 57
DB: 5 Gaps: 14

US-10-724-972A-6352 (1-335) x BW222839 (1-692)
QY 69 ValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAsnLysLys 88
Db 21 GTTCAGGATCTCTCTGTGGATGTTGCACCTTTATCTCTGGTATTGAGACC---GCT 77
QY 89 AsnArgIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArg 108
Db 78 GGAGGAGTGATGACAGCTTGG-----ATCAGAGAAACACACCGTCCCAACCAA 128
QY 109 LysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsn 128
Db 129 CAATCCCAATATTCCACCTACTTCGGCAACCAACACGAGGTGTTGATCCAGTGATAC 188
QY 129 AsnArgHisLysGlyIleTyrLysAspLeuAsnLysIle----- 141

```

```

Db 189 GAGGTGAACGATGATGACCAAGGACAAACATACCTTGGTAAAGTTCGAGCTCTCCGC 248
QY 142 -----AlaProThrIleGluLysSerPheAspGlyAspTyr 154
Db 249 ATTCCCTCCGACCTCGTGGAGTTCCAGATTGAGGTG---ACGTTGATATCGATGCT 305
QY 155 AsnGlu-----AsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGlu 172
Db 306 AACGGTATCTTTGAATGTATCTGCT-----GTTGACAAGAGCACTGGAAAGGAGAAT 356
QY 173 Glu-----GlyLysLysBargLeuGluGluHisAsp----- 182
Db 357 AAGATCAATCAACCAACGACAAAGGTGCTTTAGCAAGGAGATATTGAGCGGATGGTG 416
QY 183 LysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuPro 202
Db 417 AACGAGCCGAGAAATATAAGGAGAA-----GATGAGACGACAGAG-----GAG 461
QY 203 AlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe 222
Db 462 AAGATTCAAGCGAAGACGAGACTCGAATCTTTATGCGTTCAAC----- 503
QY 223 LeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGly---LeuSer 241
Db 504 -----CTTAAATCAACAGCTGAGGAGTGAACAAAGTCAAAAGATAAAATTTCA 548
QY 242 LysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnPro 261
Db 549 GACGAGATAAATCAGCGATTCTTAACAAAGCTAAAGAAAGTTTGGACTGGTTGGAAT 608
QY 262 GluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu 281
Db 609 AACACAG-----ACTGCGAAAGACGACGAATATGAATTCACAAAAAGAACTT 656
QY 282 GluLys-----AspProValTrpLysLysLeu 290
Db 657 GAGAAAAATTGCCAACCCCTATTATGACAAAACTT 689

RESULT 30
BW235942
LOCUS BW235942 700 bp mRNA linear EST 07-NOV-2002
DEFINITION BW235942 Nori Satoh unpublished cDNA library, tailbud embryo Ciona
intestinalis cDNA clone citb052c24 5', mRNA sequence.
ACCESSION BW235942
VERSION BW235942.1 GI:24758131
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cionidae; Ciona.
REFERENCE 1 (bases 1 to 700)
AUTHORS Satou,Y., Shin-i,T., Kohara,Y. and Satoh,N.
TITLE Expressed genes in Ciona intestinalis (2002c)
JOURNAL Unpublished (2002)
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoheascidian.zool.kyoto-u.ac.jp.

FEATURES
source
1..700
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="citb052c24"
/tissue_type="whole animal"
/dev_stage="tailbud embryo"
/clone_lib="Nori Satoh unpublished cDNA library, tailbud
embryo"

ORIGIN

```


Db 494 -----CTTAATCAACAGTGGAGGATGACAAAGTGAAGATAAAATTTCA 538
 Qy 242 LysTyLeuLysGlyProTyLeuGlnMetAsnThrGluThrLeuSerGlnValAsnPro 261
 Db 539 GACGAAGATAAATCAGGATTTCTTAACAAAGCTAAAGAAAGTTTGGACTGGTTGGAAAT 598
 Qy 262 GluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu 281
 Db 599 AACGAG-----ACTGCCGAAGAAAGGACGAATATGAATTCACAAAAGAACTT 646
 Qy 282 GluLys-----AspProValTrpLysLysLeu 290
 Db 647 GAGAAATTCGCAACCTTATTATGACAAAACCTT 679

RESULT 32

BW441129 713 bp mRNA linear EST 09-JUN-2004
 DEFINITION
 BW441129 Nori Satoh unpublished cDNA library, juvenile whole animal
 Ciona intestinalis cDNA clone cijv044124 5', mRNA sequence.

ACCESSION

BW441129

VERSION

BW441129.1 GI:48494172

KEYWORDS

EST.

SOURCE

Ciona intestinalis

ORGANISM

Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Clonidae; Ciona.

REFERENCE

1 (bases 1 to 713)

AUTHORS

Satoh, Y., Nakayama, A., Shin-i, T., Kohara, Y. and Satoh, N.

TITLE

Expressed genes in Ciona intestinalis (2004b)

JOURNAL

Unpublished (2004)

COMMENT

Contact: Nori Satoh
 Department of Zoology
 Kyoto University
 Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satoh@ascidian.zool.kyoto-u.ac.jp.
 Location/Qualifiers

FEATURES

source

1..713
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="cijv044124"
 /tissue_type="whole animal"
 /dev_stage="juvenile"
 /clone_lib="Nori Satoh unpublished cDNA library, juvenile whole animal"

ORIGIN

Alignment Scores:
 Pred. No.: 0.0275 Length: 713
 Score: 115.50 Matches: 65
 Percent Similarity: 41.04% Conservative: 38
 Best Local Similarity: 25.90% Mismatches: 91
 Query Match: 6.83% Indels: 57
 DB: 5 Gaps: 14

US-10-724-972a-6352 (1-335) x BW441129 (1-713)

Qy 69 valaspalaleuvalalaleuaspvalLysProvalGlyIleAlaAspAsnLysLys 88
 Db 34 GTTCAGGATCTTCTCTGTGGATGTTGCACCTTTATCTCTGGTATTGAGACC---GCT 90
 Qy 89 AsnArgIleIleLysProLeuArgAspLysIleGlyLysThrSerValGlyThrArg 108
 Db 91 GGAGGATGATGACAGCTTTG-----ATCAAGAGAAACACCGTCCCAACCAAA 141
 Qy 109 LysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsn 128
 Db 142 CAATCCCAATATTCACCACTTACTCGACAACCAACCGAGGTGTGTGATCCCAAGTGATC 201
 Qy 129 AsnArgHisGlyLysGlyLysAspLeuAsnLysIle----- 141

Db 202 GAGGGTGAACGTACGATGACCAAGACAAACATACCTTGGTAAGTTCGAGCTCTCCGCG 261
 Qy 142 -----AlaProThrIleGluLeuLysSerPheAspGlyAspTyr 154
 Db 262 ATCCCCCTGCACTCGTGGAGTTCACAGATTGAGGTG---ACGTTGATATCGATGCT 318
 Qy 155 AsnGlu-----AsnIleAspAlaPheLysThrIleSerLysAlaLeuLysGluGlu 172
 Db 319 AACGGTATCTTGAATGTATCTGCT-----GTTGACAAGACGACTCGAAAGGAGAAT 369
 Qy 173 Glu-----GlyLysLysArgLeuGluGluHisAsp----- 182
 Db 370 AAGATCACAATCACCACCAAGCAAGAGTCTCTTAGCAAGAGAGATATTGACGGATGGTG 429
 Qy 183 LysLysIleGluGluTyrlLysLysGluIleThrMetAspLysAsnGlnLysValLeuPro 202
 Db 430 AACGAGGCGCGAGAAATATAAGGAAGAA-----GATGAGAAGCAGAAG---GAG 474
 Qy 203 AlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe 222
 Db 475 AAGATTCAAGCGCAAGAACGACTCGAATCTTATGGCTTCAAC----- 516
 Qy 223 LeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAspValThrLysGly---LeuSer 241
 Db 517 -----CTTAATCAACGGTGGAGGATGACAAAGTGAAGATAAAATTTCA 561
 Qy 242 LysTyLeuLysGlyProTyLeuGlnMetAsnThrGluThrLeuSerGlnValAsnPro 261
 Db 562 GACGAAGATAAATCAGCGATTCTTAACAAAGCTAAAGAAAGTTTGGACTGGTTGGAAT 621
 Qy 262 GluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu 281
 Db 622 AACCGAG-----ACTCCGAAAGAGGACGAATATGAATTCACAAAAGAACTT 669
 Qy 282 GluLys-----AspProValTrpLysLysLeu 290
 Db 670 GAGAAATTCGCAACCTTATTATGACAAAACCTT 702

RESULT 33
 BW055754 753 bp mRNA linear EST 19-OCT-2002
 LOCUS
 DEFINITION
 BW055754 Nori Satoh unpublished cDNA library, blood cells Ciona intestinalis cDNA clone cibd084920 5', mRNA sequence.
 ACCESSION
 VERSION
 BW055754
 BW055754.1 GI:24156450
 EST.
 KEYWORDS
 SOURCE
 Ciona intestinalis
 ORGANISM
 Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Clonidae; Ciona.
 1 (bases 1 to 753)
 AUTHORS
 Satoh, Y., Satake, M., Azumi, K., Nonaka, M., Shin-i, T., Kohara, Y. and Satoh, N.
 TITLE
 Expressed genes in Ciona intestinalis (2002)
 JOURNAL
 Unpublished (2002)
 COMMENT
 Contact: Nori Satoh
 Department of Zoology
 Kyoto University
 Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
 Tel: 81-75-753-4081
 Fax: 81-75-705-1113
 Email: satoh@ascidian.zool.kyoto-u.ac.jp.
 Location/Qualifiers

FEATURES
 source
 1..753
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="cibd084920"
 /tissue_type="blood cells"
 /clone_lib="Nori Satoh unpublished cDNA library, blood cells"

ORIGIN

Alignment Scores:		0.0297	Length:	753
Pred. No.:	115.50	Matches:	65	
Score:	41.04%	Conservative:	38	
Percent Similarity:	25.90%	Mismatches:	91	
Best Local Similarity:	6.83%	Indels:	57	
Query Match:	5	Gaps:	14	
DB:				
US-10-724-972A-6352 (1-335) x BW055754 (1-753)				
Qy	69	ValAspAlaLeuValAlaLeuAspValLysProValGlyLeuAlaAspAsnLysLys	88	
Db	81	GTTCAGGATCTTCTGTTGGATGTCACCTTTATCTCTGTTGATTCAGACC---	GCT 137	
Qy	89	AsnArgIleLeuLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArg	108	
Db	138	GGAGGAGTGATGACAGCTTTG-----ATCAAGAGAACAACACCGTCCCAACCAAA	188	
Qy	109	LysGlnProAsnLeuGluGluIleSerLysLysProAspLeuIleAlaAspAsn	128	
Db	189	CAATCCCAATATTCACCACTTACTCGGACCAACCAAGGTGTGTTGATCCAGTGTAC	248	
Qy	129	AsnArgHisLysGlyLeuLysAspLeuAsnLysLysLysLysLysLysLysLysLys	141	
Db	249	GAGGCGTGAACGTACGATGACCAAGGACCAACATCTTGGTAACTTCGAGCTCTCCGCG	308	
Qy	142	-----AlaProThrIleGluLeuLysSerPheAspGlyAspTyr	154	
Db	309	ATCCCCCTGACCTCGTGGATTCACAGATTGAGGTG---ACCTTCGATATCATGCT	365	
Qy	155	AsnGlu-----AsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGlu	172	
Db	366	ACGGTATCTGAATGTATCTCT-----GTTGACAGACACTGGAAGGAGAT	416	
Qy	173	Glu-----GlyLysLysArgLeuGluGluHisAsp	182	
Db	417	AGATCACAAATCACCAACGACAAAGTCTCTTAGCAAGGAGATATTGAGCGAATGTG	476	
Qy	183	LysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuPro	202	
Db	477	AACGAGCGGAGAAATATAAGGAGAA-----GATGAGAACGCAAG-----GAG	521	
Qy	203	AlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe	222	
Db	522	AGATTCAGCGAAGACGGTCTCGAATCTATGCGTTCAAC-----	563	
Qy	223	LeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGly---	LeuSer 241	
Db	564	-----CTTAATCAACGGTGAAGATGACAAAGTGAAGATAAAATTTCA	608	
Qy	242	LysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnPro	261	
Db	609	GACGAAGATAAATCAGCGATTCTTAACAAAGCTAAAGAAAGTTTGGATGTTGGAAT	668	
Qy	262	GluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLysLysGluLeu	281	
Db	669	ACCAG-----ACTCCGAAAGGAGCAATATGATTTCCACAAAGAACTT	716	
Qy	282	GluLys-----AspProValTrpLysLysLeu	290	
Db	717	GAGAAATTTGCCAACCTTATTATGACAAACTT	749	
RESULT 34				
BI905625				
LOCUS				
DEFINITION				
603167626F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5255742 5',				
mRNA sequence.				
ACCESSION				
BI905625				
VERSION				
BI905625.1 GI:16168168				
KEYWORDS				
EST.				
SOURCE				
Mus musculus (house mouse)				
ORGANISM				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				

REFERENCE		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
1 (bases 1 to 981)		NIH-MGC http://mgc.nci.nih.gov/.	
AUTHORS		National Institutes of Health, Mammalian Gene Collection (MGC)	
TITLE		Unpublished (1999)	
JOURNAL		Contact: Robert Strausberg, Ph.D.	
COMMENT		Email: cgabs-r@mail.nih.gov	
		Tissue Procurement: Gilbert Smith, Ph.D.	
		cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima	
		Bonaldo, Ph.D.	
		cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
		DNA Sequencing by: Incyte Genomics, Inc.	
		Clone Distribution: NCI-CGAP clone distribution information can be	
		found through the I.M.A.G.E. Consortium/LLNL at:	
		http://image.llnl.gov	
		Plate: LLAM1646 row: c column: 07	
		High quality sequence start: 51	
		High quality sequence stop: 950.	
FEATURES		Location/Qualifiers	
source		1..981	
		/organism="Mus musculus"	
		/mol_type="mRNA"	
		/strain="CZECH II"	
		/db_xref="taxon:10090"	
		/clone="IMAGE:5255742"	
		/tissue_type="pooled lung tumors"	
		/lab_host="DH10B (phage-resistant)"	
		/clone_lib="NCI_CGAP_Lu33"	
		/note="Organ: lung; Vector: p773D-Pac (Pharmacia) with a	
		modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st	
		strand cDNA was prepared from mRNA obtained from pooled	
		lung tumors with a Not I - oligo(dT) primer [5,	
		TGTTACCACTGAAGTGGAGCGCGCTCTGTTTCTTTTCTTTT 3'].	
		Double-stranded cDNA was ligated to Eco RI adaptors	
		(Pharmacia), digested with Not I and cloned into the Not	
		I and Eco RI sites of the modified p773 vector. Library	
		went through one round of normalization, and was	
		constructed by Bento Soares and M. Fatima Bonaldo. "	
ORIGIN			
Alignment Scores:			
Pred. No.:	0.043	Length:	981
Score:	115.50	Matches:	61
Percent Similarity:	38.55%	Conservative:	40
Best Local Similarity:	23.28%	Mismatches:	106
Query Match:	6.83%	Indels:	55
DB:	4	Gaps:	9
US-10-724-972A-6352 (1-335) x BI905625 (1-981)			
Qy	87	LysLysAsnArgIleLeuLysProLeuArgAspLysIleGlyLysTyrThrSerValGly	106
Db	106	AAGGACAGCAGCTCAGCAAGCAGCAGCTGCTCCCAAAAGTACAGCAGCTGAAG	165
Qy	107	ThrArgLysGlnProAsnLeuGluGluIleSerLysLysProAspLeuIleAla	126
Db	166	AGCGAGAGAGAGAGCTGTAGTAGAAGAAAGCCAGCAGCGCTCCGAGATCTCGCAGCC	225
Qy	127	AspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGlu	146
Db	226	CAAAATCTCTGCGAAG-----CAGCGGTCTCCGTCGAGCAG	264
Qy	147	LeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLys	166
Db	265	GTGGAGGCTCGAAAAAATCTCTCAACGGCACCATTGAGCAGTGAAG-----	312
Qy	167	AlaLeuGlyLysGluGluGluGlyLysLysArg---LeuGluGluHisAspLysLysIle	185
Db	313	-----GAGGAGCTGAGGAGTAGCAGAGGTGTCTGGAGCGCGCAGCAGCAGCTG	363
Qy	186	GluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAla	205
Db	364	AGCCAGCTTCAGCAGCTGCTGAGAAATCAGAAAGAACTCTCTCGGTGACCTCGCGGAGC	423

```

QY 206 AlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyClnPheLeuSerGln 225
Db 424 CTGAAG-----CTAAAGGAAGCCCTCGAGAAAGAGGTTGGGATCATGAAGCCAGC 474
QY 226 LeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLys 245
Db 475 CTGAGGAAAGAGGAGAAAGCCAGAGAAACCAAGGAAGTCTCCANA-----525
QY 246 GlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPhe 265
Db 526 -----CTCCAGACGGAGGTTTCAGACCACCAAGCAG-----555
QY 266 IleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLysAspPro 285
Db 556 -----GCCTGAAGAACTTAGAGACCAGAGAG 582
QY 286 Val-----TrpLysLysLeuAsnAlaValLys-----294
Db 583 GTCTGTCGACATGTCGAAGTACAAAGCCACCAAGACGACTTGGAGACCAGATTTCCAAAC 642
QY 295 ----AsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArgGlyLeuLle 313
Db 643 TTAATATGACAAACTGCCAGCCTGACACGAGAGTACGACCAAGCGTGTGAGGAGAGGTC 702
QY 314 SerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAspSerLysLysAsp 333
Db 703 TCTGCCAAGGAC---GAGAAGGAGGCTGCTCCACCTGAGCATCGCAGCAGGAGATCAGGGAC 759
QY 334 AsnLys 335
Db 760 CAGAAG 765

RESULT 35
AY420059
LOCUS Mus musculus HCM7097 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY420059
VERSION AY420059.1 GI:39776016
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 4107)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 4107)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence as made by sequencing genomic exons and ordering them
based on alignment.
FEATURES
source
location
1..4107
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>4107
/locus_tag="HCM7097"
ORIGIN

```

```

Alignment Scores:
Pred. No.: 0.321 Length: 4107
Score: 115.50 Matches: 61
Percent Similarity: 38.55% Conservative: 40
Best Local Similarity: 23.28% Mismatches: 106
Query Match: 6.83% Indels: 55
DB: 9 Gaps: 9

US-10-724-972A-6352 (1-335) x AY420059 (1-4107)

QY 87 LysLysAsnArgIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGly 106
Db 3067 AAGGCAGACGAGCTCAGCAGCAGCTGAAGACCTGTCCCAAAAGTACAGCGACGTGAAG 3126
QY 107 ThrArgLysGlnProAsnLeuGluGluLysSerLysLeuLysProAspLeuLleAla 126
Db 3127 AGCCAGAGAGAGAGCTGTAGAGAAAGGCAAGCAGCGCTCCGAGATCTTCGCGAGCC 3186
QY 127 AspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGlu 146
Db 3187 CAAATCTTCTGCAGAG-----CAGCCGGTCCCGTGGAGCAG 3225
QY 147 LeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLys 166
Db 3226 GTGGAGGCTCTGAAAAAATCTCTCAACGGCACCATTGACGAGTTGAAG-----3273
QY 167 AlaLeuGlyLysGluGluGluGlyLysLysArg---LeuGluGluHisAspLysLysIle 185
Db 3274 -----GAGGAGCTGAGGAGTAAAGCAGAGGTCTCTGGAGCGGAGCAGCAGCGGTG 3324
QY 186 GluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAla 205
Db 3325 AGCCAGCTTTCAGCAGCTCTCGAGAAATCAGAAGAACTCTCTGGTACCTCGCGGAGCAGC 3384
QY 206 AlalysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyClnPheLeuSerGln 225
Db 3385 CTGAAG-----CTAAGGAAGCCCTCGAGAAAGAGTTGGGATCATGAAGCCAGC 3435
QY 226 LeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLysSerLysTyrLeuLys 245
Db 3436 CTGAGAGAAAGGAGAGAGAAAGCCAGAAAGAAACCAAGGAAGTCTCCAAA-----3486
QY 246 GlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPhe 265
Db 3487 -----CTCCAGACGAGGTTTCAGACCACCAAGCAG-----3516
QY 266 IleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLysAspPro 285
Db 3517 -----GGCTGAGAACTTAGAGACCAGAGAG 3543
QY 286 Val-----TrpLysLysLeuAsnAlaValLys-----294
Db 3544 GTGTCGACATGTCGAGAGTACAAAGCCACCAAGACGACTTGGAGACCAGATTTCCAAC 3603
QY 295 ----AsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArgGlyLeuLle 313
Db 3604 TTAATATGACAAACTCGCCAGCCTGAACAGGAAGTACGACCAAGCGCTGTGAGGAGAGGTC 3663
QY 314 SerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAspSerLysLysAsp 333
Db 3664 TCTGCCAAGGAC---GAGAAGGAGCTGCTCCACCTGAGCATCGCAGCAGGAGATCAGGGAC 3720
QY 334 AsnLys 335
Db 3721 CAGAAG 3726

RESULT 36
CR426562
LOCUS CR426562 XGC-tailbud Xenopus tropicalis cdna clone TTBA033p07 5,
DEFINITION mRNA sequence.
ACCESSION CR426562
VERSION CR426562.1 GI:48919971

```

```

KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Xenopus tropicalis (western clawed frog)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 785)
Croning, M.D.R., Ashurst, J.L., Taylor, R., Garrett, N. and Rogers, J.
Sanger Xenopus tropicalis EST project 2001 (2004)
TROPICALIS_SEQUENCE_ID: TBA033p07.p1kSP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Nigel Garrett.
Seq primer: SP6.

FEATURES
    Location/Qualifiers
        1..785
            /organism="Xenopus tropicalis"
            /mol_type="mRNA"
            /db_xref="taxon:8364"
            /clone="TBA033p07"
            /dev_stage="tailbud (stage 28-30)"
            /lab_host="Escherichia coli DH10B."
            /clone_lib="XGC-tailbud"
            /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
            was oligo dt primed from Sug of poly A+ RNA from tailbud.
            EcoRI-NotI cut cDNA was then ligated into pCS107 with
            EcoRI at the 5' end and NotI at the 3' end."

ORIGIN
Alignment Scores:
Pred. No.:      0.0358      Length:      785
Score:          115.00      Matches:    59
Percent Similarity: 37.2%      Conservative: 27
Best Local Similarity: 25.54%      Mismatches: 99
Query Match:      6.80%      Indels:     46
DB:               7          Gaps:         6

US-10-724-972A-6352 (1-335) x CR426562 (1-785)
Qy 76 AspValLysProValGlyLeuAlaAspAsnLysLysAsnArgIleIleLysProLeu 95
Db 91 GATGTGAAGAGGTTAAAGTTCTGTGTAATTCGCAAACTTACTCGACATCGCAAAAGCTG 150
Qy 96 ArgAspLysIleGlyLysTyrThrSerVal-----GlyThrArgLysGlnProAsn 112
Db 151 AAATCTAAAGGGGACAGAGAACTAATCTTAAGATCAGAGGAGCCGACAGATTG 210
Qy 113 LeuGluGluIle-----SerLysLysLysProAspLeuIleIleAlaAspAsn 129
Db 211 CTTGAGGAATCCATCTCTGTTAAGGAGCTCAACACAGATGATGTACTAAAACTGCTTTA 270
Qy 130 ArgHisLysGlyIleTyr-----Lys 136
Db 271 AGGAAGAAATCAGTTTGAATAATATGCAAAAAGCTTAACAGCACTGCTGAGGAGCGA 330
Qy 137 AspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGlu 156
Db 331 GCATTAGCAAGACTGGCCACACACCTTTGCTTTAAGCAGAAATACACAGCTATTAAAGAG 390
Qy 157 AsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLys---GluGluGluGlyLys 175
Db 391 GCTATTAAAGCTTTTAAAGATGCAAGAAAACTGCTGCTGAAGGGGAGGGAAGAGAG 450
Qy 176 LysArgLeuGluGluHisAspLysIleGluGluTyrLysLysGluIleThrMetAsp 195
Db 451 AAGGATGAACCTGACAGGTCAAGAAGATTAAAGAAACCAAAAACCTGTTACAAACGAAG 510
Qy 196 LysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSer 215

```

```

Db 511 CTAAATRAAATAACAGAAGAGATATAAATCACAAAAGAACATGTTAAAGAGAGAAATGC 570
Qy 216 AsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAsp 235
Db 571 AAAAAC-----CTGCTAGAAAGAT 588
Qy 236 ValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThr 255
Db 589 TCAGATAGGGCACTGAAAAGCTCTTGAACCTATTATGTGCAAGAGAACTTGCTGCTGAG 648
Qy 256 LeuSerGlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSerAsnGlu 275
Db 649 CAGACAGCAGAAAACAAGGA-----CAG 672
Qy 276 ProSerLeuLysGluLeuGluLysAspProVal 286
Db 673 CCGAAGGCCCAAGATGTGGAAAGGCCACCTGTG 705

RESULT 37
LOCUS CR731171
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR731171
VERSION CR731171.1 GI:51229441
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE
1 (bases 1 to 1120)
Genoscope.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
: 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : segre@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
    source
        1..1120
            /organism="Tetraodon nigroviridis"
            /mol_type="mRNA"
            /db_xref="taxon:99883"
            /tissue_type="fish"

ORIGIN
Alignment Scores:
Pred. No.:      0.059      Length:      1120
Score:          115.00      Matches:    60
Percent Similarity: 40.48%      Conservative: 59
Best Local Similarity: 20.41%      Mismatches: 99
Query Match:      6.80%      Indels:     76
DB:               3          Gaps:         11

US-10-724-972A-6352 (1-335) x CR731171 (1-1120)
Qy 78 LysProValGlyIleAlaAspAsnLysLys----- 88
Db 162 CGACCGAGCGACACGCGAGCGGCGCAAGAAAGCGCGCGAGGACAGTGCACACAGCTG 221
Qy 89 -----AsnArgIleIleLysProLeuArgAspLysIleGlyLys 101
Db 222 GAGGAGGAGCTGTGGTGTCTCGAAGAAAGCTGAAGGGAGTGCAGAGTGCAGCTGACAG 281
Qy 102 TyrThrSerValGlyThrArgLysGlnProAsnLeuGluIleSerLysLysLysPro 121
Db 282 TACTCCGAGTCCCTGAAGGACGCCCGAGAGAGCTGGAGAGCGGAGGAGGAGGAGGAGCA 341
Qy 122 AspLeu-----IleIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeu 138

```

```
Db 342 GATCGGAGCGGAGCTGGCTCTCTGAACCGCGGATCCAGCTGCTGGAGGAGGCTG 401
Qy 139 AsnLysIleAlaProThrIle-----GluLeuLysSerPheAspGlyAsp 153
Db 402 GACCGGCGCCAGGAGCGACTGGCCACCGCCCTCCAGAGCTGGAGGAGCGCGAGAGGCT 461
Qy 154 TyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu--- 172
Db 462 GCACGACGAGCGAGCGAGAGGAATGAAGGTGATTGAGAACAGACGACGCAAAAGACGAGGAG 521
Qy 173 -----GluGlyLysLysArgLeuGluGluHisAsp 182
Db 522 AAGATCGAGATCCAGGAGATGAGCTGAAGGAGGCCAAGCACATCCCGAGGAGGCCGAC 581
Qy 183 LysLysIleGluGluTyrLysLysGluIleThr-----MetAspLysAsn 197
Db 582 CGCAAGTACGAGGAGGTGGCTCGAAACTCGGTATCCTGGAAGGAGACCTGGAGCGCTCC 641
Qy 198 GlnLysValLeuProAlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSer 217
Db 642 GAGGAACGCGCGAGGTGGCGGAGGCCAAATCTGGAGATCTT----- 683
Qy 218 TyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAspValThr 237
Db 684 ----- 595
Qy 238 LysGlyLeuSerLysTyrLysLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSer 257
Db 696 AAAAAGCTCACCAACATCTTGAAGCTCCGGAAGCTCAG-----GCCGAGAAGTACTCC 749
Qy 258 GlnValAsn-----ProGluArgMetPheIleMetThrAsnLysAlaSerSerAsn 274
Db 750 CAAAAAGAGGACAAGTATGAAGGAGGATCCGAGTCTCTGACGAGCAAG----- 797
Qy 275 GluProSerLeuLysGluLeuGlyAspProValTyrLysLysLeuAsnAlaValLys 294
Db 798 -----CTGAGGAGCGCGAGACTCGTGGGAGTTTCTGAGAGGTCGGTGCCCAAG 848
Qy 295 AsnGlnArg---ValAspIleLeuAspArgAspLeuTyrAlaArgSer---ArgGlyLeu 312
Db 849 CTGGAGAAGACCATCAGCATCTAGAAGACGAGGCTCTACGCTCAGAAAGCTGAAGGCGCAAG 908
Qy 313 IleSerSerGluMetAlaLysGluLeuValGluLeuSer 326
Db 909 GCTCTGAGCGAGGAGCTGGACCTGGCCCTGAATGACATGACC 950

RESULT 38
LOCUS CR685229 2252 bp mRNA linear HTC 19-AUG-2004
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR685229
VERSION CR685229.1 GI:51183136
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
1 (bases 1 to 2252)
Genoscope.
Direct Submission
Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
Submitted (10-AUG-2004) CP 5706 - 91057 EVRY cedex - FRANCE.
(S-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
Location/Qualifiers
1..2252
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
```

```
ORIGIN
Alignment Scores:
Pred. No.: 0.157 Length: 2252
Score: 115.00 Matches: 70
Percent Similarity: 29.06% Conservative: 55
Best Local Similarity: 21.88% Mismatches: 103
Query Match: 6.80% Indels: 92
DB: 3 Gaps: 16

US-10-724-972A-6352 (1-335) x CR685229 (1-2252)
Qy 34 SerSerLysGluSerSerLysAspGlyValGluLysHisGlu-----Glu 49
Db 1068 TCTCAAAAGAGCGCCAAAGATGGATAAAGTCAAGT---CACGATATCGTGTGTCGCGC 1124
Qy 50 GlyThrThrLysValProLysHisProLysArgVal----- 61
Db 1125 GGTTCACCTCGCATCCCAAAATCCAGAAACTGCTCCAGGACTACTTCAACGGGAGGTTT 1184
Qy 61 ----- 61
Db 1185 TGAACAAGAGCATCAACCTGATGAAGCTGTTGCTTATGTCGCCGCTGTCAGGAGGCCA 1244
Qy 62 -----ValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysPro 79
Db 1245 TCTTGTCAAGTGCACAAAGTCTGAGAATTGTGCAGGACCTGCTCTCTCGACGCTCACTCT 1304
Qy 80 ValGlyIleAlaAspAspAsnLysLysAsnArgIleLysProLeuArgAspLysIle 99
Db 1305 CTGTCTCTTGTGATTGAGAGC---GCCGAGGCGTCATGACCGTGCTG-----ATC 1352
Qy 100 GlyLysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluLysSerLysLeu 119
Db 1353 AAGCGCAACACCACTCTTACCAAGCAGACGCCAGACCTTCACCACTACTCTGACACAC 1412
Qy 120 LysProAspLeuIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsn 139
Db 1413 CAGCCGCGCGTCTCATCAGGTTTACGAGGGTGCGCTGTCATGACGAGGACCAACAT 1472
Qy 140 LysIle-----AlaProThrIle 145
Db 1473 CTGTCGGGCAAGTTCGAGTTGACGGGCATCCCCCTGCGCCCTGTCGGGCTTCCCAGATC 1532
Qy 146 GluLeuLysSerPheAspGlyAspTyrAsn-----GluAsnIleAspAlaPheLysThr 163
Db 1533 GAGGTG---ACTTTTGACATCGACGCCCAACGGCATCATGACGTGCTGCC----- 1580
Qy 164 IleSerLysAlaLeuGlyLysGluGlu-----GlyLysLysArgLeu 178
Db 1581 GTGGACAAAAGCACCTGGCAAGAAAACAAGATCACCATCACCACCAATGCAAAAGTCTCTC 1640
Qy 179 GluGluHisAsp-----LysLysIleGluTyrLysLysGluIleThr 193
Db 1641 AGTAAGGAGGACATTGAGCGCATGTGCGAGGAAGTGAAGATCAAGCCGACGACGAC 1700
Qy 194 MetAspLysAsnGlnLysValLeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHis 213
Db 1701 GTCCAGCGTGACAAG-----GTGTCGGCCCAAGATGGCTTGTGAGTCGTAC 1745
Qy 214 ProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSer 233
Db 1746 GCTTTCAC--ATGAAGTCAACCGTGGAA 1772
Qy 234 AspAspValThrLysGly---LeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsn 252
Db 1773 GACGAGAACTCGCTGGGAAGATCAGTGCAGATGACAGACAGAGATTTTGGACAGTGC 1832
Qy 253 ThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSer 272
Db 1833 AACGAGGTCATCAGCTGGCTGGCAAGAAATCAG-----ACTGCCGAGGAGGAC 1880
```

/tissue_type="Eyes"

QY 273 SerAenGluProSerLeuLysGluLeuGluLys-----AspProValTrpLysLysLeu 290
 Db 1881 GAGTACGAGCATCAGCAGAGGAGGTGGAGAGGTTGCAACCCCATCATCACAAAGATG 1940

RESULT 39

LOCUS BJ706325 707 bp mRNA linear EST 08-MAR-2004
 DEFINITION BJ706325 MF01FFA cDNA Oryzias latipes cDNA clone MF01FFA017e14 5',
 mRNA sequence.

ACCESSION BJ706325

VERSION BJ706325.1 GI:45247204

KEYWORDS Oryzias latipes (Japanese medaka)

SOURCE Oryzias latipes

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
 Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.

REFERENCE 1 (bases 1 to 707)
 AUTHORS Kohara,Y., Shin-i,T., Kimura,T., Marita,T., Jindo,T. and Takeda,H.
 TITLE Medaka EST Project in Takeda's lab
 JOURNAL Unpublished (2001)
 COMMENT Contact: Tadaou Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.

FEATURES
 Location/Qualifiers
 1..707
 /organism="Oryzias latipes"
 /mol_type="mRNA"
 /strain="HG-rr"
 /db_xref="taxon:8090"
 /clone="MF01FFA017e14"
 /sex="mixture of female and male"
 /tissue_type="whole embryo"
 /dev_stage="fry stage 40"
 /clone_lib="MF01FFA cDNA"

ORIGIN
 Alignment Scores:
 Pred. No.: 0.0352 Length: 707
 Score: 114.50 Matches: 61
 Percent Similarity: 42.23% Conservative: 45
 Best Local Similarity: 24.30% Mismatches: 88
 Query Match: 6.78% Indels: 57
 DB: 4 Gaps: 13

US-10-724-972A-6352 (1-335) x BJ706325 (1-707)

QY 69 ValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLysLys 88
 Db 12 GTCCAAAGACTTCTGCTTCTGGATGCTCACTCCACTTCTCTGGGTATAGAGACT---GCT 68

QY 89 AsnArgIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArg 108
 Db 69 GGTGGGTGATGACAGTCTTA-----ATCAACGCAATACTACAAATTCCTACCAAA 119

QY 109 LysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsn 128
 Db 120 CAACCCAGACTTTACTATACACTCTGTATACAGCTGGGGTCTATCCAGTTTAT 179

QY 129 AsnArgHieLysGlyIleTyrLysAspLeuAsnLysIle----- 141
 Db 180 GAAGGTGAGCGTCCATGACCAAGCAACACCTGCTGGGAAAGTTTGAACGTGACAGA 239

QY 142 -----AlaProThrIleGluLeuLysSerPheAspGlyAspTyr 154
 Db 240 ATTCTCTCTGCTCCTCGTGGTGTCTCTCAGATTGAGGTG---ACATTGTATATTCACGA 296

QY 155 Aen-----GluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGlu 172

Db 297 AATGAATCATGAATGTTTCTGCT-----OTTGATAAGACACTGGCAAGGAAAC 347
 QY 173 Glu-----GlyLysLysArgLeuGluGluHisAsp----- 182
 Db 348 AAAATCAACAATCACTAATAACAAAGGTCGCTTTAGTAAAGAGGACATTTGAGCGCATGGTC 407

QY 183 LysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuPro 202
 Db 408 CAAGAAGCTGAGAAGTACAAGGCAAGATGATGTTTCAGCGTGACAA----- 455

QY 203 AlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe 222
 Db 456 ---GTGCTGCAAGAATGGCTGAGTCTTATCTTCAAC----- 494

QY 223 LeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGly---LeuSer 241
 Db 495 -----ATGAAGTCAACTGTGGAGATGAGAAGCTTGTCTGGCAAGATCAGT 539

QY 242 LysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnPro 261
 Db 540 GATGAGGACAAACAAAGATCTTGGATAAGTGAATGAATGATCAGCTGGCTGGACAG 599

QY 262 GluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeu 281
 Db 600 AACACAG-----ACTGCTGAAGAGATGAGTACGAGCACCAACAGAGGAGCTT 647

QY 282 GluLys-----AspProValTrpLysLysLeu 290
 Db 648 GAGAAAGTATGTAATCCCATCATCACCAAACTG 680

RESULT 40
 LOCUS AL666132 711 bp mRNA linear EST 11-JAN-2002
 DEFINITION AL666132 directional larval cDNA library Ciona intestinalis cDNA
 clone 002ZE10 5', mRNA sequence.

ACCESSION AL666132
 VERSION AL666132.1 GI:18133039
 KEYWORDS EST.
 SOURCE Ciona intestinalis
 ORGANISM Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 Phlebobranchia; Cionidae; Ciona.

REFERENCE 1 (bases 1 to 711)
 AUTHORS Ciona intestinalis directional larval cDNA library
 JOURNAL Unpublished (2002)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 IMPORTANT: this sequence may contain errors. The Ciona intestinalis
 library from which the clone was isolated may be contaminated with
 cDNAs from bacteria or other Eukarya.

Directional larval cDNA library originate from Dr.M.Branno,
 Stazione A.Dohrn, Naples, Italy, and was prepared in
 pBluescript2SK+.

FEATURES
 Location/Qualifiers
 1..711
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="002ZE10"
 /clone_lib="directional larval cDNA library"
 /note="Vector: pBluescript2SK+."

ORIGIN
 Alignment Scores:
 Pred. No.: 0.0355 Length: 711
 Score: 114.50 Matches: 67
 Percent Similarity: 41.67% Conservative: 38
 Best Local Similarity: 26.59% Mismatches: 89
 Query Match: 6.78% Indels: 58

DB:	1	Gaps:	14
US-10-724-972A-6352 (1-335) x AL666132 (1-711)			
Qy	69	ValAspAlaLeuValalaLeuAspVallyysProvalGlyIleAlaAspAsnLysLys	88
Dd	19	GTTCAGGATCTTCCTTGGATGTGCACCTTATCTCTGGATTGAGACC---	75
Qy	89	AsnArgllelleLysProLeuArgAspLyslleGlyLysTySerValGlyThrArg	108
Dd	76	GGAGGAGTGATCACAGCTTTG-----ATCAAGAGAACAACGACTGTCCCAACCAA	126
Qy	109	LysGlnProAsnLeuGluIleSerLysLysLeuYsProAspLeuIlelealaspAsn	128
Dd	127	CAATCCCNAATATTCCACCATTTACTCGGAACCAACCAAGGTGTGTGTATCCAAGTGTA	186
Qy	129	AsnArgHisLysGlyIleTyLysAspLeuAsnLyslle-----	141
Dd	187	GAGGTTGAACGTACGATGACCAAGCAACAACATACTTGGTAAGTTCGAGCTCTCCGCG	246
Qy	142	-----AlaProThrIleGluLeuLysSerPheAspGlyAspTyr	154
Dd	247	ATCCCCCTGCACCTCGTGGAGTTCCACAGATTGAGTG---ACGTTGCATATCATGTCT	303
Qy	155	AsnGlu-----AsnIleAspAlaphelysThrIleSerLysAlaleuGlyLysGluGlu	172
Dd	304	AACGGTATCTTGAATGTATCTGCT-----GTTGACAAGCACTGGAAGGAGAA	354
Qy	173	Glu-----GlyLysLysArgLeuGluGluHisAsp-----	182
Dd	355	AAGATCACAATCACCAACGACANAGGTCTGTACCAAGAAGATATTGAGCGGATGGT	414
Qy	183	LysLyslleGluGluTyLysGluIleThrMetAspLysAsnGlnLysValLeuPro	202
Dd	415	AACGAGGCCGAGAAATAAAGAACAA-----GATGACAAGCAGAA--GAG	459
Qy	203	AlaValAlaAlaLysSerGlyLeuLeualahisProSerAnsSerTyValGlyGlnPhe	222
Dd	460	AAGATTCAAGCAAGACCGACTCGAATCTTATGCTTCAAC-----	501
Qy	223	LeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGly---LeuSer	241
Dd	502	-----CTTAAATCAACGGTGGAGGATGACAAGTCAAGATCAAGATTCA	546
Qy	242	LysTyrlsLysGlyProTyrlsGlnMetAnthrGluThrLeuSerGlnValAsnPro	261
Dd	547	GACGAAGATAAATCAGCGATTCTTTAACAAAGCTAAAGAAGTTTTTAGACTGGTTGAAA	606
Qy	262	Glu-ArgMetPheIleMetThrAsnLysAlaserSerAsnGluProSerLeuLysGluLe	281
Dd	607	AACCAAGCTGCGGAAAGGACGATATCAATTCACAA-----AAAGAACT	654
Qy	281	uGluLys-----AspProValTrpLysLysLeu	290
Dd	655	TGAGAAAAATTCGCAACCCCTATATGACAAACTT	688

Search completed: November 10, 2005, 00:11:33
Job time : 3894 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 9, 2005, 19:57:58 ; Search time 935 Seconds
(without alignments)
2962.983 Million cell updates/sec

Title: US-10-724-972A-6352
Perfect score: 1690
Sequence: 1 GVESVRLKLTSLVIGLLFVL.....EEMAKELVELSKDKKDKK 335

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delxet 7.0

Searched: 9794790 seqs, 4134909567 residues

Total number of hits satisfying chosen parameters: 19589580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool_h/US10724972/runat_03112005_141902_29179/app_query.fasta_1.519
-DB=Published Applications NA -QWMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10724972 @CGN 1.1.480 @runat_03112005_141902_29179
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOC=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELETE=7

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
- 23: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 24: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
- 26: /cgn2_6/ptodata/1/pubpna/US11A_NEW_PUB.seq.*
- 27: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1690	100.0	1008	24	US-10-724-972A-2580	Sequence 2580, Ap
2	1671	98.9	993	18	US-10-282-122A-34954	Sequence 34954, A
3	1242	73.5	1014	9	US-09-815-242-8156	Sequence 8156, Ap
4	1232	72.9	984	18	US-10-282-122A-7644	Sequence 7644, Ap
5	1181	69.9	927	9	US-09-815-242-4404	Sequence 4404, Ap
6	726	43.0	957	9	US-09-815-242-4194	Sequence 4194, Ap
7	726	43.0	960	9	US-09-815-242-8024	Sequence 8024, Ap
8	724.5	42.9	957	22	US-10-470-048B-301	Sequence 301, Ap
9	724.5	42.9	960	18	US-10-282-122A-7949	Sequence 7949, Ap
10	644	38.1	525	18	US-10-282-122A-35107	Sequence 35107, A
11	516.5	30.6	897	18	US-10-282-122A-30706	Sequence 30706, A
12	510.5	30.2	801	9	US-09-974-300-2066	Sequence 2066, Ap
13	475	28.1	668	8	US-08-781-986A-355	Sequence 355, Ap
14	475	28.1	668	19	US-10-329-624-355	Sequence 355, Ap
15	460	27.2	900	18	US-10-282-122A-19870	Sequence 19870, A
16	460	27.2	909	9	US-09-815-242-6372	Sequence 6372, Ap
17	460	27.2	909	18	US-10-282-122A-20621	Sequence 20621, A
18	460	27.2	10244	22	US-10-893-671-14	Sequence 14, Appl
19	454	26.9	891	18	US-10-282-122A-33236	Sequence 33236, A
20	405	24.0	990	22	US-10-470-048B-318	Sequence 318, Ap
21	405	24.0	999	17	US-10-278-946-15	Sequence 15, Appl
22	405	24.0	999	22	US-10-967-189-15	Sequence 15, Appl
23	405	24.0	3775	8	US-08-781-986A-238	Sequence 238, Ap
24	405	24.0	3775	19	US-10-329-624-238	Sequence 238, Ap
25	391	23.1	972	18	US-10-282-122A-9415	Sequence 9415, Ap
26	380.5	22.5	944	9	US-09-974-300-2061	Sequence 2061, Ap
27	375	22.2	2209	24	US-10-450-763-7183	Sequence 7183, Ap
28	375	22.2	4392	24	US-10-450-763-8842	Sequence 8842, Ap
29	350	20.7	2799	24	US-10-450-763-13139	Sequence 13139, A
30	334	19.8	2115	8	US-08-781-986A-604	Sequence 604, Ap
31	334	19.8	2115	19	US-10-329-624-604	Sequence 604, Ap
32	333	19.7	242	8	US-08-781-986A-2556	Sequence 2556, Ap
33	333	19.7	242	19	US-10-329-624-2556	Sequence 2556, Ap
34	331	19.6	945	9	US-09-738-626-375	Sequence 375, Ap
35	331	19.6	3309400	9	US-09-738-626-1	Sequence 1, Appli
36	329	19.5	648	9	US-09-974-300-2087	Sequence 2087, Ap
37	327	19.3	963	18	US-10-282-122A-17083	Sequence 17083, A
38	326	19.3	915	24	US-10-721-922A-351	Sequence 351, Ap
39	321	19.0	957	9	US-09-974-300-2044	Sequence 2044, Ap
40	317.5	18.8	321	9	US-09-815-242-2107	Sequence 2107, Ap
41	317.5	18.8	321	18	US-10-282-122A-4605	Sequence 4605, Ap
42	313	18.5	942	18	US-10-282-122A-10099	Sequence 10099, A
43	301	17.8	942	18	US-10-282-122A-9201	Sequence 9201, Ap
44	300.5	17.8	708	9	US-09-974-300-2068	Sequence 2068, Ap
45	296.5	17.5	1056	24	US-10-724-972A-3409	Sequence 3409, Ap

ALIGNMENTS

RESULT 1

US-10-724-972A-2580
; Sequence 2580, Application US/10724972A
; Publication No. US2004014734A1
; GENERAL INFORMATION:
; APPLICANT: Doucette-Stamm, Lynn
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE OF INVENTION: PATH03-16
; FILE REFERENCE: PATH03-16
; CURRENT APPLICATION NUMBER: US/10/724, 972A
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: 05/450, 969
; PRIOR FILING DATE: 1999-11-29

721 AGTAAAGTATCTTAAAGGACCTTACTTACAAATGAACACTGAAACCTTTATCTCAAGTGAAT 780
261 ProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGlu 280
781 CTGAGCGTATGTTTCAATATGACAAACAAAGCAAGTCTTAACGACACCTTCACTTAAAGAA 840
281 LeuGluLysAspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIle 300
841 CTAGAAAAAGATCCTCTGATGGAAGAAATTAACGCTGTGAAAAATCAACGCTGTGTATTT 900
301 LeuAspArgAspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLys 320
901 TTAGACCGTGACTTATGGGCAAGATCACGTGGTTAATTTCTTCAGAGAAGAAATGGCAAAA 960
321 GluLeuValGluLeuSerLysLysAspSerLysLysAspAsnLys 335
961 GAACTGTGTAATTAATCTAAGAAAGATAGTAAAAAAGATAATAAG 1005

RESULT 2
US-10-724-972A-34954
; Sequence 34954, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A.
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34954
; LENGTH: 993
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-10-724-972A-34954

Alignment Scores:
Pred. No.: 5,83e-134 Length: 993
Score: 1671.00 Matches: 331
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

1 GlyValGluSerValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeu 20
1 GGAGTGAATCAGTGAAGAGGTTTAAAAATTTAAGTGAATTTGGCTTATTTGTTTGTAA 60
21 IleAlaThrAlaAlaCysGlyAsnAsnSerSerSerAsnSerLysGluSerLys 40
61 ATTGCAACTGCGCATGTGGAAATAATAGTTCAAGTAACCTCAAGTAAAGAGTCATCAAAA 120
41 AspGlyValGluIleLysHisGluGluGlyThrThrLysValProLysHisProLysArg 60
121 GATGAGTTGAAATCAAGCACGAAGAGGTACTACGAAAGTACCTAAACACCCCTAAACGT 180
61 ValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProVal 80
181 GTTGTGTTCTTGAGTATTCATTTTGTGATGCTGTAGTTGCTTTTGAATGTTTAACTGTT 240
81 GlyIleAlaAspAspAsnLysLysAsnArgIleLysProLeuArgAspLysIleGly 100
241 GGGATAGCGGATGATACAAAAAATCGTATTATTAACCATTAAGAGATAAAATCGA 300
101 LysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluLysSerLysLys 120
301 AAATACACTTCTGTAGGAACACGTAAAGCAACCTTAACCTTAGAAGAATCAGTAACTTAA 360
121 ProAspLeuIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLys 140
361 CCAGATTTTAAATTTATGCTGATTAATATAGACAAAGGTATTATTAAGACTTAAATAA 420
141 IleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAla 160
421 ATTGCTCCCTACCATTTGAATGAAAGTTCGATGGAGATTATATGAAATAATTTGATGCT 480
161 PheLysThrIleSerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGlu 180
481 TTTAAACAATTTCAAAAGCTTTAGCTAAAGAAGAAGAGGTAAAGGCTTTAGAGAA 540
181 HisAspLysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysVal 200
541 CACGATAAGAAAAATTTGAAGAATAATAAAAAAGAAATAATATGATGATTAATAATCAAAAGGTA 600
201 LeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGly 220
601 TTGCTCCGAGTAGCTGCTTAATCAGGTTTGTCTGCTCATCCCAAGCAACTCTTATGTTGGT 660
221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspAspValThrLysGlyLeu 240
661 CAACTTCTAAGTCAACTAGGTTTTTAAAGAAGCATTAAGTATGATGTTACTAAGGTTTA 720
241 SerLysTyrLeuLysGlyProTyrLeuGluMetAsnThrGluThrLeuSerGlnValAsn 260

```

RESULT 3
US-09-815-242-8156
; Sequence 8156, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haeselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8156
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1014)
US-09-815-242-8156

Alignment Scores:
Pred. No.: 4,24e-97 Length: 1014
Score: 1242.00 Matches: 239
Percent Similarity: 84.89% Conservative: 42
Best Local Similarity: 72.21% Mismatches: 50
Query Match: 73.49% Indels: 1
DB: Gaps: 0

US-10-724-972A-6352 (1-335) x US-09-815-242-8156 (1-1014)
QY 1 GlyValGluSerValArgGlyLeuIysIleLeuSerValIleGlyLV
Db 20 GGAGTGGGAACG-ATGAGAGGCTCTAAAACCTTTAGTATATTGGGAT
QY 21 IleAlaThrAlaAlaCysGlyAsnAsnSerSerSerAsnSerSerL
Db 79 TTTTGTAGTTGCAGCTGTGGTAATACGGATAATTCAAGTAAAAAAG
QY 41 AspGlyValGluIleIysHisGluGlyThrThrIysValProL
Db 139 GATACTATTTCGTTAAAGATGAAANTGGTACGTAAGTAAGTACCTTA
QY 61 ValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuA
Db 199 ATCGTTGTATTAGACTACTCTATTTCGCAGATGCATTAGCAGCATTA
QY 81 GlyIleAlaAspAsnIysIysAsnArgIleIleIysProLeuA
Db 259 GGTATTTCGTGATGATGGTAAGAAAAACGTTATCATTTAAACACGATTGA

```

101 LysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluLeuSerLysLeuLys 120
Db |||||
319 AATTATACCTCTGTAGTACACGTAACAGCCAACTTAGAGGAATTTAGTAATTAATA 378
Qy |||||
121 ProAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 140
Db |||||
379 CCGATTTAATATTCGCTGATAGCAGTAGACATAAAGGTATTAATAAGAAATTAACAAA 438
Qy |||||
141 IleAlaProThrLeuGluLeuLysSerPheAspGlyAspTyrAsnGluLeuLeuLeu 160
Db |||||
439 ATTGCACCAACATTTATCAAGAGTTTGTGAGGAGCTACAAACAAACATTAATTCG 498
Qy |||||
161 PheLysThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 180
Db |||||
499 TTCAAAACAATGCTAAAGCTTTAAATAAGAAAGAAAGCGGCGGCGCTTCCTGCTGAA 558
Qy |||||
181 HisAspLysLysLeuGluGluTyrLysLysGluLeuLeuLeuLeuLeuLeuLeuVal 200
Db |||||
559 CATGATAAATTAATCAAAAGATATAAGATGAAATTAAGTTTGATAGAAATCAAAAGTG 618
Qy |||||
201 LeuProAlaValAlaAlaLysSerGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuValGly 220
Db |||||
619 CTTCCAGCAGTTGTTGCTAAAGCTGTTATTAAGCACAATCCAAACTATTCATATGTTGGA 678
Qy |||||
221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu 240
Db |||||
679 CAATTTTAAACGAACTTGGAATTTAAATGCAATTAAGTATGATGTACAAAGGTTTA 738
Qy |||||
241 SerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsn 260
Db |||||
739 AGTAAATACCTTGAAGGACCTTACTTACAATAGTACTGAACATTTAGCTGACTTAAT 798
Qy |||||
261 ProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGlu 280
Db |||||
799 CCTGAACGATGATTTATATGACAGATAATGCTAAAAAGATTCCTGCAATTCGAAGAAG 858
Qy |||||
281 LeuGluLysAspProValTyrLysLysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 300
Db |||||
859 TTACAAGAAGATCCAACTTGAAAGATGTAACGCGAGTTAAATAATTCGGTGGATATT 918
Qy |||||
301 LeuAspArgAspLeuTyrAlaArgSerArgGlyLeuLeuSerSerGluGluMetAlaLys 320
Db |||||
919 GTTGACCGTGATGTTTGGCAAGATCTCGTGGCTTAATTTCTTCTGAAGAATGCTAAA 978
Qy |||||
321 GluLeuValGluLeuSerLysLysAspSerLys 331
Db |||||
979 GAACCTGTTGATTAATCAAAAAAGAACAAAG 1011

RESULT 4
US-10-282-122A-7644
; Sequence 7644, Application US/10282122A
; Publication NO. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIORITY APPLICATION NUMBER: 60/191,078
; PRIORITY FILING DATE: 2000-03-21
; PRIORITY APPLICATION NUMBER: 60/206,848
; PRIORITY FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patentin version 3.1
SEQ ID NO 7644
LENGTH: 984
TYPE: DNA
ORGANISM: Staphylococcus aureus
US-10-282-122A-7644

Alignment Scores:
Pred. No.: 2,95e-96 Length: 984
Score: 1232.00 Matches: 235
Percent Similarity: 84.71% Conservative: 42
Best Local Similarity: 71.87% Mismatches: 50
Query Match: 72.90% Indels: 0
Gaps: 18

US-10-724-972A-6352 (1-335) x US-10-282-122A-7644 (1-984)

Qy 5 ValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeuLeuAlaThrAla 24
Db |||||
1 ATGAGAGCTCTAAACAACTTTTAGTATATTTGGATTAATAGTTCCTTACTTTTGTGCTGCA 60
Qy 25 AlaCyGlyAsnAsnSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 44
Db |||||
61 GCTTGTGTGTAATACGGATAATTCAGTAAAAAGAAAGATCATCACTAAAGATATTTTCG 120
Qy 45 IleLysHisGluGluGlyThrLysValProLysHisProLysArgValValValLeu 64
Db |||||
121 GTAAAGATGAAATGGTACAGTAAAGTACATAAGATGCAAAACGATATCGTTGTATTA 180
Qy 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyLysAlaAsp 84
Db |||||
181 GAGTACTCATTTGCAGATGCATTAGCAGCTTTAGACGTTAAACCCAGTTGGTATTCGTGAT 240
Qy 85 AspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSer 104
Db |||||
241 GATGTAAGAAAAACGATATCTTAAACCGATTAGAGAAAAAATTTGGGATTTACTTCT 300
Qy 105 ValGlyThrArgLysGlnProAsnLeuGluLeuLeuSerLysLeuLysProAspLeuLeu 124
Db |||||
301 GTAGTACAGCTTAAACAGCAAACTTAGAGGAAATTAGTAAATTAACCCGATTTAAAT 360
Qy 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144
Db |||||
361 ATCGCTGATAGCAGTAGACATTAAGGTATTAATAAGAAATTAACAAAAATTCACCAACA 420
Qy 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164
Db |||||
421 TTATCATTAAGAGTTTTGATGGAGACTACAAACAAATATTAATTCGTTCAAAACAAT 480
Qy 165 SerLysAlaLeuGlyLysGluGluGlyLysArgLeuGluGluLysAspLysLys 184
Db |||||
481 GCTAAGCTTTAAATAAGAAAAAGGCGAAAAACGCTTGTGCTGAGCATGATAAATTA 540
Qy 185 IleGluGluTyrLysLysGluLeuThrMetAspLysAsnGlnLysValLeuProAlaVal 204
Db |||||

370 AAAATCAATCAGAAATTTGAGCAAAATTTGCTCCGACAATCATGTTAGTTCGGGTACGGGA 429
153 AspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGluGlyLysGluGlu 172
430 GATTATAATCAATATTTGAAGCATTTAAACAGTCGCTTAAGCAGTAGGCAAGAGAAA 489
173 GluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGluTyrLysLysGluIle 192
490 GAAGGCGAGAGAGCGTCTGGAAAGCATGATATAATATTAGCGGAGATTAGAAAGAAAT 549
193 ThrMetAspLysAsnGlnLysValLeuProAlaValAlaAlaLysSerGlyLeuLeuAla 212
550 GAACAGAGTACGTTAAATCTGCAATTTGCAATTCGCGGTATCTCAAGAGCAGGTATGTTATT 609
213 HisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232
610 AATAATGAAGATACATTTATGGGACAAATCTTAAATTAATTAATTAATTAATTAATTAAT 669
233 SerAspAspValThrLysGlyLeuSerLysTyrLysGlyLeuSerLysTyrLysGlyLeu 252
670 ACAAAAGACAAACTACGCATGTTGGTGAACGCGCAAGGTTGCTCTTATATATATATTAAT 729
253 ThrGluThrLysSerGlnValAsnProGluArgMetPheIleMetThrAsn---LysAla 271
730 AATGAAGAACTTGCAATATCAATCCAAAGATTATGATTTTAGCCACTGACGCAAAACG 789
272 SerSerAsnGluProSerLeuLysGluLeuGluLysAspPro---ValTyrLysLysLeu 290
790 GACAAAAATAGACCAAAATTC-----ATTGATCTCGCAGTTTGGAAATCATTA 837
291 AsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTyrPalaArgSerArg 310
838 AAAGCTGGAAGAGATACAAAGTTTATGACGTTGACCGAATAAGTGGTTGAATCAAGG 897
311 GlyLeuLysSerSerGluGluMetAlaLysGluLeuValGluLeuLeuSerLysLys 328
898 GGGATTATCGCAAGTGAAGTATGCGCAGAGATTTAGAAAAAATTTGCAGAAAAA 951

US-10-724-972A-7949
; Sequence 7949, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-06-28
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 2000-03-47
; PRIOR FILING DATE: 2000-09-09
; PRIOR FILING DATE: 2000-04-27
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27

Alignment Scores: 1,12e-52 Length: 960
Pred. No.: 724,50 Matches: 149
Score: 66.67% Conservative: 63
Percent Similarity: 46.86% Mismatches: 95
Best Local Similarity: 42.87% Indels: 11
Query Match: 18 Gaps: 4
DB: 1

US-10-724-972A-6352 (1-335) x US-10-282-122A-7949 (1-960)
Qy 13 ValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCysGlyAsnAsnSerSerSer 32
Db 25 GTTGTGTTTATGCTAAATCTTAGTTGTAGCAGTAGCGGGTTGTGCTCAAAAGATACT --- 81
Qy 33 AsnSerSerLysGluSerSerLysAspGlyValGluIleLysHisGluGluGlyThr 52
Db 82 -----GAAGAGAAAACTGNAATGACGCAATAAAAGATGAATTTAGGAACCTGAA 129
Qy 53 LysValProLysHisProLysArgValValValLeuGluTyrSerPheValAspAlaLeu 72
Db 130 AAAATTAAGAAAAATCTTAACGTTGTTGTATTAGATATAGTTTGTGCTGATTTATTA 189
Qy 73 ValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLysLysAsnArgIleIle 92
Db 190 GCAGCATTTAGATATGAAACCTGTTGGTATTGCAGATGATGCGCAGCTTAAATAATAACA 249
Qy 93 LysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsn 112
Db 250 AAGTCAGTAAGAGATAAGATTGGGGCATATGAATCGGTTGGATCTAGACCGCAACCGAAT 309
Qy 113 LeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsnAsnArgHisLys 132
Db 310 ATGGAAGTGAATGAATTAATTAACCGGATTTGATCATTCAGATGTTAGCAGACATAAG 369
Qy 133 GlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGly 152
Db 370 AAAATCAATCAGAAATTCAGCAAAATTCCTCCGACAATCATGTTAGTTCGGGTACGGGA 429
Qy 153 AspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 172
Db 430 GATTATAATCAATATTTGAAGCATTTAAACAGTCGCTTAAGCAGTAGGCAAGAGAAA 489
Qy 173 GluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGluTyrLysLysGluIle 192
Db 490 GAAGGCGAGAGAGCGTCTGGAAAGCATGATATAATATTAGCGGAGATTAGAAAGAAAT 549
Qy 193 ThrMetAspLysAsnGlnLysValLeuProAlaValAlaAlaLysSerGlyLeuLeuAla 212
Db 550 GAACAGAGTACGTTAAATCTGCAATTTGCAATTCGCGGTATCTCAAGAGCAGGTATGTTATT 609
Qy 213 HisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeu 232
Db 610 AATAATGAAGATACATTTATGGGACAAATCTTAAATTAATTAATTAATTAATTAATTAAT 669
Qy 233 SerAspAspValThrLysGlyLeuSerLysTyrLysGlyLeuSerLysTyrLysGlyLeu 252
Db 670 ACAAAAGACAAACTACGCATGTTGGTGAACGCGCAAGGTTGCTCTTATATATATTAAT 729

Qy 253 ThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn---LysAla 271
Db 730 AATGAAGACTTGGCAATATCAATCCAAAGATTATGATTTAGCCACTGACGGAACG 789
Qy 272 SerSerGlnGluProSerLeuLysGluLeuGluLysAspPro---ValTrpLysLysLeu 290
Db 790 GACAAAATAGAACGAATTC-----ATTGATCTCGAGTTTGGAAATCATTA 837
Qy 291 AsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSerArg 310
Db 838 AAAGCTGTGAAGAGTAACAAAGTTTATGACGTTGACCGGAATAAGTGGTTGAATCAAGG 897
Qy 311 GlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328
Db 898 GGGATTATCCAGTGAAGATGATGGCAGAGATTAGAAAAAATTGCAAAAAA 951

RESULT 10

US-10-282-122A-35107
; Sequence 35107, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35107
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Staphylococcus haemolyticus
US-10-282-122A-35107

Alignment Scores:

Pred. No.:	4.41e-46	Length:	525
Score:	644.00	Matches:	124
Percent Similarity:	84.00%	Conservative:	23
Best Local Similarity:	70.86%	Mismatches:	24
Query Match:	38.11%	Indels:	4
DB:	18	Gaps:	1

US-10-724-972A-6352 (1-335) x US-10-282-122A-35107 (1-525)

Qy 5 ValArgGlyLeuLysIleLeuSerValIleGlyLeuLeuPheValLeuIleAlaThrAla 24
Db 1 ATGAAGGCTTTAAGTTTGCAGGTATAGTCGCATATATTATTGTCATTAGTTTACT 60
Qy 25 AlaCysGlyAsnAsnSerSerSerSerSerSer-----SerLysGluSerSerLys 40
Db 61 GCTTGTGGTAAATGTATAGCAATAATGGATCAGGTGATTCAGGTAATAAATCGTCATCGAA 120
Qy 41 AspGlyValGluIleLysHisGluGluGlyThrThrLysValProLysHisProLysArg 60
Db 121 GATTCATTATAAATTAAACACGAATTAGGCACAACTAAAGTTCTTAAAGATGCTAAACGT 180
Qy 61 ValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProVal 80
Db 181 GTCGTAGCGTTGGAATTTTCATTTGTAGATGCTTTAGCGCAGCTAAATGTTAAACCTGTT 240
Qy 81 GlyIleAlaAspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGly 100
Db 241 GGGGTTGCTGATGACAAACAAACCAATCGTATTATTAAACCACTTAAAGAAAAATTGGA 300
Qy 101 LysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluLysSerLysLeuLys 120
Db 301 GATTATAAATCTGTGTGCTCTGTAACCAACCACTTAGAAGAAATCAGTAAATTAATA 360
Qy 121 ProAspLeuIleIleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLys 140
Db 361 CCAGATTATAATTATTGTCAGATAGTAACACACACAAAGGTATTTCACAAAGATTAGATAA 420
Qy 141 IleAlaProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAla 160
Db 421 ATTGCTCCAACTTGAATTAATAAAGTTTGTATGAGATTATAATGATAATATAGATGCT 480
Qy 161 PheLysThrIleSerLysAlaLeuGlyLysGluGluGluGlyLys 175
Db 481 TTTAAACAAATGCCCAGCTTTAAATAAAGATGATGTAGGTGAC 525

RESULT 11

US-10-282-122A-30706
; Sequence 30706, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining prior application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30706
; LENGTH: 897
; TYPE: DNA
; ORGANISM: Pasteurella multocida
US-10-724-972a-6352 (1-335) x US-10-282-122A-30706 (1-897)
Alignment Scores:
Pred. No.: 7,61e-35 Length: 897
Score: 516.50 Matches: 111
Percent Similarity: 60.48% Conservative: 65
Best Local Similarity: 38.14% Mismatches: 98
Query Match: 30.56% Indels: 17
DB: 4 Gaps: 4
US-10-724-972A-6352 (1-335) x US-10-282-122A-30706 (1-897)
Qy 43 ValGluLeuLysHisGluGluGlyThrThrLysValProLysHisProLysArgValVal 62
Db 67 GTTACCGTCAAGATCAAAAGGGGAATTTACCTAGATAGCGTACCTAAACGTTGTT 126
Qy 63 ValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyLe 82
Db 127 GCCTTAGAATATTCTTATGTGGATCGACTTCACAAATTTGGTGTCAGCCCTTGGTGTC 186
Qy 83 AlaAspAsnLysLysAsnArgLysLysProLysProLysValProLysHisGlyLysTyr 102
Db 187 GCAGATGCAATATATAAACCCGATTTCTGCAAAAGTACCGATTAAGTCGCGCATGG 246
Qy 103 ThrSerValGlyThrArgLysGlnProAsnLeuGluGluLeuSerLysLysProAsp 122
Db 247 GAATCAGTGGGAGCGCGTTCTCAACGAGTTTGAAGCGATTTCTGCACTTAACCCAGAT 306
Qy 123 LeuLeuLeuAlaAspAsnAsnArgHisLysGlyLysLysLysAspLeuAsnLysLeuAla 142
Db 307 TTGATCAATTCGCGATGATAATCGCCATTTCTGCGGCTATGAAGAACTCAAAAAATTCGG 366
Qy 143 ProThrLeuGluLysSerPheAspGlyAspTyrAsnGluAsnLeuAspAlaPheLys 162
Db 367 CCGACAGTCTGCTTTAATTTCCCGCCATGAGAACTATCAAGAAACCTTGAACCCGACAA 426
Qy 163 ThrLeuSerLysAlaLeuGlyLysGluGluGlyLysLysLysLysLysLysLysLysLys 182
Db 427 AAATTCGGTGATTATTAGTAAATCAAAAGAAATGCAAGCGCTATTGCAAAACATAAG 486
Qy 183 LysLysLeuGluGluTyrLysLysGluLysLysLysLysLysLysLysLysLysLysLys 202
Db 487 CAGGATATTTCGACATCGCCAAA-----ACGTTACCGAAAGGGGAAAGCGATTATC 540
Qy 203 AlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe 222
Db 541 GGTGTTTTCAGTGAACCCAAATTTAATTTATATATATAGCAATCTTATGCTGGTGTCTTA 600
Qy 223 LeuSerGlnLeuGlyPheLys-----GluAlaLeuSerAspAspValThrLysGlyLeu 240
Db 601 GTGGAAGTGTAGTTATCAATTCGCAAAAGCCGCGTCCGATAAC----- 645
Qy 241 SerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrClnuThrLeuSerGlnValAsn 260
Db 646 -----CAACCTAATGCTTCGTTGTTGTTAGAACCAAGTGGCGGCGAGAAAG 690
Qy 261 ProGluArgMetPheLeuMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGlu 280
Db 691 CCTGATCTGATGATCTTAATCCATTATTCGT-----GATGAGAGTATTGCAAGAAA 741

281 LeuGluLysAspProValTrpLysLysLysLysLysLysLysLysLysLysLysLysLys 300
Db 742 TGGGAAATGAAGCCTTATGGAAATATATCCCTGGGTAAAAAATGGTCAAGTCAATCTTA 801
Qy 301 LeuAspArgAspLeuTrpAlaArgSerArgGlyLeuLeuSerSerGluGluMetAlaLys 320
Db 802 GCTAATGATAATTTATGGCAAGACGCTGTTGATGGCCCTGAAGTAATGGCTTAA 861
Qy 321 GluLeuValGluLeuSerLysLysLysLysLysLysLysLysLysLysLysLysLys 331
Db 862 GAAGTCCAAGACTTTGTAAAGAAATCCGCCAAA 894
RESULT 12
US-09-974-300-2066
; Sequence 2066, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2066
; LENGTH: 801
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-2066
Alignment Scores:
Pred. No.: 2,17e-34 Length: 801
Score: 510.50 Matches: 108
Percent Similarity: 62.88% Conservative: 58
Best Local Similarity: 40.91% Mismatches: 95
Query Match: 30.21% Indels: 3
DB: 9 Gaps: 3
US-10-724-972A-6352 (1-335) x US-09-974-300-2066 (1-801)
Qy 60 ArgValValValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysPro 79
Db 3 AAAGTTGTGGTCTTTGAATTTGGGATTTTATGATGGCTCTTGTATGGCGGTATTAAAGCT 62
Qy 80 ValGlyLysLeuAlaAspAsnLysLysLysLysLysLysLysLysLysLysLysLysLys 99
Db 63 GTCCGAATTCGCGACGCGCGGCAAACTAAGTTTATTAACGAGAAGGTTCAGGGGAAAAATC 122
Qy 100 GlyLysTyrThrSerValGlyThrArgLysGlnProAsnLeuGluGluLysLysLysLeu 119
Db 123 AAAGGTATATCTTCAGTCGGTTCGCGCCGCGCCAGCAAGCTTTGAAAAAATTCCTCTTTA 182
Qy 120 LysProAspLeuLeuLeuAlaAspAsnAsnArgHisLysGlyLysLysLysLysLysLys 139
Db 183 AAGCCGATTTAATTTATTTCCGATTCGACGACGCGCGGTCTATGATAGCTGTGCG 242
Qy 140 LysLeuAlaProThrLeuGluLysSerPheAspGlyAspTyrAsnGluAsnLeuAsp 159
Db 243 AAAATTCGCCCGCAATATCGCTCAAAATACTTGAATCCGATTTATCAGGACACGCTTGAT 302
Qy 160 AlaPheLysThrLysSerLysAlaLeuGlyLysGluGluGluGluLysLysLysLysLys 179
Db 303 GCATCTCTTACGATTCGAAAGCGCTCGCAAGAAAGCAATGAGAAAAAATTTGGCT 362
Qy 180 GluHisAspLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 199
Db 180 Lys 199

Db 363 GAACATAAACAAAGCTGGATGAAATGAAACAGAAATTCGGC---AGCCGGAACAGAGC 419
QY 200 ValLeuProAlaValAlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrVal 219
Db 420 ATTCTTCTGCTCGGGAACACAAATGAGAAATCACCGTGCAGATGAAACATTTTCACG 479
QY 220 GlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGly 239
Db 480 TCTCAGCTTTCGAGAAATCGGTATACATACCGTGTGCGAGACAGC---GGCAAGGC 536
QY 240 LeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnVal 259
Db 537 GATGCCGAAAGCGTGAATCGTCAATATTAAATGACGCTCGAACAGCTGCTTGAGAAA 596
QY 260 AsnProGluArgMetPheLeuMetThr---AsnLysAlaSerSerAsnGluProSerLeu 278
Db 597 GATCCCCGACGTTATTGTCTCTGATGACGCGGAGAGAAAGATAAAGTCGATGAGACGGAAA 656
QY 279 LysGluLeuLysAspProValTrrpLysLeuAsnAlaValLysAsnGlnArgVal 298
Db 657 AGACCATCGAAAGAGATCTCTTTGGAAAAAGCTCAGCGCAGTCAAAAAACGGCAAGGTT 716
QY 299 AspIleLeuAspArgAspLeuTrrpAlaArgSerArgGlyLeuIleSerSerGluGluMet 318
Db 717 TATGAAGCCGACAGATTCGCTCGCTCGGACGAGCATTCAGCGAGCAGATGAGCTG 776
QY 319 AlaLysGluLeu 322
Db 777 ATGATGAATC 788
RESULT 13
US-08-781-986A-355
; Sequence 355, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 355:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 668 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-355
Alignment Scores:

Pred. No.: 1.95e-31 Length: 668
Score: 475.00 Matches: 91
Percent Similarity: 86.89% Conservative: 15
Best Local Similarity: 74.59% Mismatches: 16
Query Match: 28.11% Indels: 0
DB: 8 Gaps: 0
US-10-724-972A-6352 (1-335) x US-08-781-986A-355 (1-668)
QY 210 LeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLys 229
Db 2 TTTATTTAGCACATCCAACTATTTCATATGTTGGACAATTTTAAACGAACATAGGATTTAAA 61
QY 230 GluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeu 249
Db 62 AATGCATTAAGTGACCATGTAAACAAAGGTTTAAGTAATATTGAAAGGACCTTACTTA 121
QY 250 GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheIleMetThrAsn 269
Db 122 CAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGCGTATGATCATATTGACAGAT 181
QY 270 LysAlaSerSerAsnGluProSerLeuLysGluLeuLysAspProValTrrpLysLys 289
Db 182 CATGCTAAAAAGATTTCTGCTGAATTTCAAGAGCTTACAAGACATGCAACATGGAAGAA 241
QY 290 LeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrrpAlaArgSer 309
Db 242 TTGATGCGAGTTAAAAAATAATCGCGTGGATATTGTTGACCGTGATGTTGGCAAGATCT 301
QY 310 ArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAsp 329
Db 302 CGTGCGTTAATTTCTTCTGAGAGAAATGCGCTAAAGAACTTGTGTAATTTATCAAAAAAGAA 361
QY 330 SerLys 331
Db 362 CAAAAG 367
RESULT 14
US-10-329-624-355
; Sequence 355, Application US/10329624
; Publication No. US20040043037A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/329,624
; FILING DATE: 27-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/956,171
; FILING DATE: October 20, 1997
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:

```

; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248PID1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 355:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 668 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 355:
US-10-329-624-355

Alignment Scores:
Pred. No.: 1,95e-31 Length: 668
Score: 475.00 Matches: 91
Percent Similarity: 86.89% Conservative: 15
Best Local Similarity: 74.59% Mismatches: 16
Query Match: 28.11% Indels: 0
DB: 19 Gaps: 0

US-10-724-972A-6352 (1-335) x US-10-329-624-355 (1-668)
QY 210 LeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLys 229
Db 2 TTATTAGCACATCAAACTATTTCATATGTTGGCAATTTTAAACGAACACTAGGATTATAA 61
QY 230 GluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeu 249
Db 62 AATGCATTAAGTCAGCATGTAACAAAGGTTTAACTAATAATTTGAAAGGACCTTACTTA 121
QY 250 GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMetPheLeuMetThrAsn 269
Db 122 CAATTAGACACTGAACATTTAGCTGATTTAAATCCAGAGCGTATGATCATTTATGACAGAT 181
QY 270 LysAlaSerSerAsnGluProSerLeuLysGluLeuGluLysAspProValTyrLysLys 289
Db 182 CATGCTAAAAAGATTTCTGCTGAATTCAGAAGTTTACAAGAGATGCAACATGGAAAG 241
QY 290 LeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTyrAlaArgSer 309
Db 242 TTGAATGCAGTTAAAAAATAATCGCGTGATATTGTTGACCGTGATGTTGGCAAGATCT 301
QY 310 ArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAsp 329
Db 302 CGTGCTTAATTTCTTCTGAAGAAATGCTTAAGAACTTGTGTAATTATCAAAAAAGAA 361
QY 330 SerLys 331
Db 362 CAARAG 367

RESULT 15
US-10-282-122A-19870
; Sequence 19870, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19870
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Enterobacter cloacae
US-10-282-122A-19870

Alignment Scores:
Pred. No.: 5,45e-30 Length: 900
Score: 460.00 Matches: 101
Percent Similarity: 57.09% Conservative: 60
Best Local Similarity: 35.82% Mismatches: 105
Query Match: 27.22% Indels: 16
DB: 18 Gaps: 4

US-10-724-972A-6352 (1-335) x US-10-282-122A-19870 (1-900)
QY 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64
Db 70 GTTCAAGCAGCAACACGCGACGCTTTACACTCGATAAACCCACACGCGATTGTGTGCTG 129
QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyLeAlaAsp 84
Db 130 GAATCTCGTTCGCGGATCGCTGCGCGCGTGCAGCGTGCAGCGGATCGGTATTGCCGAC 189
QY 85 AspAsnLysLysAsnArgIleIleLysProLeuArgAspLysLleGlyLysTyrThrSer 104
Db 190 GATAACGATGCACAAACGATCTGCGCGAAGTGCCTGCGCACCTGAAACCCGTGGCAGTCC 249
QY 105 ValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIle 124
Db 250 GTCGGAACGCGCGCGCAGCGCGAGCTGGAAGCCATTTGTCCTCGAACCAGACCTGATC 309
QY 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144
Db 310 ATTCGCGACAGCAGTCGCGCATCGCGGGATTACACTGCTTCGACAAATCGCGCGGTA 369
QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164
Db 370 CTGCTCTTAAGTCCCGCAACGAAACCTACGCTGAAAAATTTGCACTCAGCGGCTATCATC 429
QY 165 SerLysAlaLeuGlyLysGluGluGlyLysLysLysLysLysLysLysLysLysLys 184
Db 430 GCGCAAGTGGTAGGTAATAAAACGAGAGATGCGGCGGCTGTGGAACAACATAAAGAGAAG 489
QY 185 IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204
Db 490 ATGGCGCATGTGTCCAGCCAG-----CTTCCCACAGGACACGCGCTTGGCACA 543
QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224
```


Db 544 TCACGCGAAGACAGTATTTCACCTGCATACCCAGAGACCTGGAGCCGGAGCGTCTGGCT 603
QY 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu 244
Db 604 TCTCTGGGGCTG-----AACGTTCCCGCTGCGATG 633
QY 245 LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263
Db 634 GCGGCGCGTCCATGCGCTCCATCGGCTCGAGCAACTGCTGGCGGTCAATCCTCGCTGG 693
QY 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLys 283
Db 694 CTGCTGGTTCACCACTATCG-----GAAGAGAGCATTTGTTAAACGCTGCGCAACA 744
QY 284 AspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg 303
Db 745 GATCCGCTCTGGCAGATATTAACCGCGCGCAGAGCAGCAGGTGCTTCAGTCGACAGT 804
QY 304 AspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuVal 323
Db 805 AACGCTGGCGCGGATGCGCGTATTTTTCGCGCAGAGCGTATTGCGCGTCGACAGGTA 864
QY 324 GluLeu 325
Db 865 AAAATC 870

RESULT 16

US-09-815-242-6372.
; Sequence 6372, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIORITY FILING DATE: 2000-03-21
; PRIORITY FILING DATE: 2000-03-21
; PRIORITY FILING DATE: 2000-05-23
; PRIORITY FILING DATE: 2000-05-23
; PRIORITY FILING DATE: 2000-05-26
; PRIORITY FILING DATE: 2000-10-23
; PRIORITY FILING DATE: 2000-10-23
; PRIORITY FILING DATE: 2000-11-27
; PRIORITY FILING DATE: 2000-12-22
; PRIORITY FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6372
; LENGTH: 909
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(909)
US-09-815-242-6372

Alignment Scores: 5.52e-30 Length: 909
Pred. No.: 460.00 Matches: 100
Score:

Percent Similarity: 57.45% Conservative: 62
Best Local Similarity: 35.46% Mismatches: 104
Query Match: 27.22% Indels: 16
DB: 9 Gaps: 4
US-10-724-972A-6352 (1-335) x US-09-815-242-6372 (1-909)
QY 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64
Db 76 GTTCAGAGCAACACGCGCGTTTACACTCGAAAAACGCCAACCGATTGGTGGCTG 135
QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84
Db 136 GAACTCTCGTTGCGCGATGCGTGGCGCGTGGAGTGGTATTCGCGATTCGATTCGCGAC 195
QY 85 AspAsnLysLysAsnArgIleLysProLeuArgAspLysIleGlyLysTyrThrSer 104
Db 196 GATAACGATGCAAAACGCGATCTGCGCGAAGTGGTGGCGCACCTGAAACCGTGCAGTCC 255
QY 105 ValGlyThrArgLysGlnProAsnLeuGluLysSerLysLysProAspLeuIle 124
Db 256 GTCCGAAACGCGCGCGAGCGCGAGCTTGGAAAGCCATTGCGCGCTCTGAAACGACACCTGATC 315
QY 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144
Db 316 ATTGCGCAGCAGCAGTGGCGGTTTACATCGCTTGCAGCAATTCGCGCGGTA 375
QY 145 IleGluLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrile 164
Db 376 CTGCTGCTTAAAGTCCCGCAACGAAACCTACGCTGAAATTTGCAATCTCGCGGTATCATC 435
QY 165 SerLysAlaLeuGlyLysGluGlyLysLysArgLysLysLysLysLysLysLys 184
Db 436 GCGCAAAATGGTGGTAAAGCAGAGATGACGAGGACGCTCTGGAACAAACATAAGAGAGG 495
QY 185 IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204
Db 496 ATGGCGCAGTGGCGCCAGCCAG-----CTTCCCAAGGGACACGCGTGGCCTTTGGCACA 549
QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224
Db 550 TCACGCGAAGACAGCAATTCAACCTGCATCTACGAGAGACCTGACCGGACGCTGCTGGCC 609
QY 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLysSerLysTyrLeu 244
Db 610 TCTCTGGGGCTG-----AACGTTCCCGCTGCGATG 639
QY 245 LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263
Db 640 GCGGTGGCTCCATGCCCTGCGCTGAGCACTGCTGGCGGTCAATCCTCGCTGG 699
QY 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLysLysGluLeuGluLys 283
Db 700 CTGCTGGTTGCCCACTATCGC-----GAAGAGAGCATTTGTTAAACGCTGGCAACA 750
QY 284 AspProValTrpLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg 303
Db 751 GATCCGCTCTGGCAGATGTTAAACCGCGCGCAGAGCAGCAGGTTCGTTTCGCTGACAGT 810
QY 304 AspLeuTrpAlaArgSerArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuVal 323
Db 811 AACACCTGGCGCGGATGCGCGTATTTTTCGCGAGCGGTATTGCGCGCTGACACCGTA 870
QY 324 GluLeu 325
Db 871 AAAATC 876

RESULT 17

US-10-282-122A-20621
; Sequence 20621, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 20621
LENGTH: 909
TYPE: DNA
ORGANISM: Escherichia coli
US-10-282-122A-20621

Alignment Scores:
Pred. No.: 5,52e-30 Length: 909
Score: 460.00 Matches: 100
Percent Similarity: 57.45% Conservative: 62
Best Local Similarity: 35.46% Mismatches: 104
Query Match: 27.22% Indels: 16
DB: 18 Gaps: 4

US-10-724-972A-6352 (1-335) x US-10-282-122A-20621 (1-909)

QY 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64
Db 76 GTTCAGGACGACACGCGACGTTTACATCGAAAAACCCACACGGAATGTGTGTG 135
QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84
Db 136 GAATCTCTGTCGCGATCGGTGGCGCGGTGAGCGTATCCGATCGGTATTCGCGAC 195
QY 85 AspAsnLysLysAsnArgIleLysProLysArgAspLysLysLysLysLysLysLys 104
Db 196 GATAACGATGCAAAACGATCTGCGCGAGTGGCGCGACCTGAAACCGTGGCGAGTCC 255
QY 105 ValGlyThrArgLysGlnProLeuGluGluLysSerLysLysLysProAspLeuIle 124
Db 256 GTTCGGAACGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 315
QY 125 IleAlaAspAsnArgHisLysGlyIleTyrLysAspLeuAsnLysLysLysLysLys 144
Db 316 ATTCGCGACGACGATCGCCATGCGGGGGTTTACATCGCTTGCAGCAAAATCGCGCGGTA 375

QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164
Db 376 CTGCTGCTTAAGTCCCGCAACGAACTACGCTGAAAAATTTGCAATCTGGCGGTATCATC 435
QY 165 SerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluHisAspLysLys 184
Db 436 GCGCAAAATGTTGGTAAAGCGAGATGCGACGCGTCTGGAACAACATAAAGAGAGG 495
QY 185 IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204
Db 496 ATGGCGAGTGGCGCCAGCCAG-----CTTCCCAAGGGACACGCGTGGCGCTTTGGCACA 549
QY 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPheLeuSer 224
Db 550 TCACGCAACAGCAATTTCAACCTGCATCTCGGCTGAGCAACTGCTGGCGGTCAATCTCTGCGCC 609
QY 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLysLysLysLysLys 244
Db 610 TCTCTGGGCGTG-----AACGTTCCCGTGGCGATG 639
QY 245 LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263
Db 640 GCGGTCGCTCCATGCGCTCATCGGCTGAGCAACTGCTGGCGGTCAATCTCTGCGCTGG 699
QY 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLysLys 283
Db 700 CTGCTGTTGCCACTATCG-----GAAGAGAGCATTTGTTAAACGCTGGCAACAA 750
QY 284 AspProValTyrLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg 303
Db 751 GATCCGCTCTGGCAGATGTTAACCGCGCGCGAGCAGCAGGTTGCTTCGTCGACAGT 810
QY 304 AspleuTyrAlaArgSerArgGlyLeuLysSerGluGluMetAlaLysGluLeuVal 323
Db 811 AACACCTGGCGCGAGTGGCGGTATTTTCTGCGAGCGGTATTCGCGCTGACACGTA 870
QY 324 GluLeu 325
Db 871 AAAATC 876

RESULT 18
US-10-893-671-14/c
; Sequence 14, Application US/10893671
; Publication No. US20050064527A1
; GENERAL INFORMATION:
; APPLICANT: Levy, Stuart, et. al
; TITLE OF INVENTION: NMR COMPOSITIONS AND THEIR METHODS OF USE
; FILE REFERENCE: PKZ-043
; CURRENT APPLICATION NUMBER: US/10/893,671
; CURRENT FILING DATE: 2004-07-15
; PRIOR APPLICATION NUMBER: US/09/801,563
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/188,362
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 10244
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-893-671-14

Alignment Scores:
Pred. No.: 1.07e-28 Length: 10244
Score: 460.00 Matches: 100
Percent Similarity: 57.45% Conservative: 62
Best Local Similarity: 35.46% Mismatches: 104
Query Match: 27.22% Indels: 16
DB: 22 Gaps: 4

US-10-724-972A-6352 (1-335) x US-10-893-671-14 (1-10244)

```
Qy 45 IleYshisGluGluGlyThrLysValProLysHisProLysArgValValValLeu 64
Db 7735 GTTCAGGACGACGACGACGTTTACACTGAAAACGACGACGATGTTGGTCTG 7676
Qy 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84
Db 7675 GAACCTCTGTTGGCGATGCGTGGCGCGTGGACGTCATCCGATCGGTATGCCGAC 7616
Qy 85 AspAsnLysLysAsnArgIleLysPheProLeuArgAspLysLysGlyLysTyrThrSer 104
Db 7615 GATAACGATGCAAAACGATCCTCCGCGAAGTCGTCGCGACCTGAAACGTCGCGAGTCC 7556
Qy 105 ValGlyThrArgLysGlnProAsnLeuGluLysSerLysLysLysProAspLeuLeu 124
Db 7555 GTCCGAACGCGCGGACGCGGAGCTTGAAGCATTGCCCTCTGAACACGACCTGATC 7496
Qy 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144
Db 7495 ATTGCGGACAGCAGTGCCTGCGGGGTTTACATCGCTTGCAGCAATCGCGCGGTA 7436
Qy 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164
Db 7435 CTGCTCTTAAGTCCCGCAGAACTACGCTGAAATTTGCAATCTGCGGCTATCATC 7376
Qy 165 SerLysAlaLeuLysGluGluGlyLysLysArgLeuGluLysHisAspLysLys 184
Db 7375 GCGCAATGTTGGTAAAGCGAGATGCGAGCAGCTGCGAAACATAAGAGAGG 7316
Qy 185 IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204
Db 7315 ATGGCGCAGTGGGCCAGCCAG-----CTTCCAAAGGACGACGCTGGCCCTTGGCACA 7262
Qy 205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlnPheLeuSer 224
Db 7261 TCACGCGAAGCAGCAATTCAACCTGCATCTCAGGAGACCTGGACCGGCGCTGTGGCC 7202
Qy 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu 244
Db 7201 TCTCTGGGCTG-----AACGTTCCCGCTGCGATG 7172
Qy 245 LysGlyProTyrLeu---GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg 263
Db 7171 GCGGGTGCCTCCATCGCGTCCATCGCGCTGGAGCACTGCTGGCGGTCAATCTGCTGG 7112
Qy 264 MetPheIleMetThrAsnLysAlaSerSerAsnGluProSerLysGlyLeuLys 283
Db 7111 CTGCTGGTGGCCACTATCGC-----GAAGAGAGCATTTGTTAAACGCTGGCAACA 7061
Qy 284 AspProValTyrLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArg 303
Db 7060 GATCCGCTCTGGCAGATGTTAAACCGCGCAGAGACGAGGTTGCTTGGTTCGACAGT 7001
Qy 304 AspLeuTyrAlaArgSerArgGlyLeuLysSerGluLysSerGluLysAlaLysLeuVal 323
Db 7000 AACACCTGGCGCGGATGCGCGGTATTTTGTGCGAGAGCGTATTTGCGCGCTGACACGTA 6941
Qy 324 GluLeu 325
Db 6940 AAAATC 6935

RESULT 19
US-10-282-122A-33236
; Sequence 33236, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Lianguo
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33236
; LENGTH: 891
; TYPE: DNA
; ORGANISM: Pseudomonas syringae
US-10-282-122A-33236

Alignment Scores:
Pred. No.: 1,76e-29 Length: 891
Score: 454.00 Matches: 110
Percent Similarity: 52.80% Conservative: 60
Best Local Similarity: 34.16% Mismatches: 104
Query Match: 26.86% Indels: 48
DB: 18 Gaps: 7

US-10-724-972A-6352 (1-335) x US-10-282-122A-33236 (1-891)
Qy 17 LeuPheValLeuIleAlaThrAlaAlaCysGlyAsnAsnSerSerSerAsnSerLys 36
Db 13 CTGCTCACGCTGCTGGCGAGCGCGCT----- 39
Qy 37 GluSerSerLysAspGlyValGluLysHisGluGluGlyThrLysVal----- 54
Db 40 -----CAGGCTGCGCCCATCGACATCGATGACGCGCACGACCAAGTGTCATCTG 87
Qy 55 ProLysHisProLysArgValValValValLysSerPheValAspAlaLeuValAla 74
Db 88 CCAGACACCCCAAGCGCGTGGTACTGTAATTCGTAATTCGTAATTCGTAATTCG 147
Qy 75 LeuAspValLysProValGlyIleAlaAspAsnLysLysAsnArgIleLysPro 94
Db 148 GTTGGCGTGACACCGTGGCGCGCGCAGATGATGGCGACGCGCAGTGTGTTACTGCCAAG 207
Qy 95 LeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnProAsnLeuGlu 114
Db 208 GTGCGCAAGGCGGTGGGTGTAATGTCAGTGGTGGGCTGCGCTCGCAACCAATATCGAA 267
Qy 115 GluLysSerLysLeuLysProAspLeuIleAlaAspAsnArgHisLysGlyIle 134
Db 268 GTGATCGCAGCGCTCAAGCGCGACCTGATCATCGCTGACCTCGGTGCTGTCATCGCTCT 327
Qy 135 TyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGlyAspTyr 154
Db 135 TyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAspGlyAspTyr 154
```


Alignment Scores:	
Pred. No.:	3,286-25
Score:	405.00
Percent Similarity:	51.74%
Best Local Similarity:	32.85%
Query Match:	23.96%
DB:	22
US-10-724-972A-6352 (1-335) x US-10-967-189-15 (1-999)	
Length:	999
Matches:	113
Conservative:	65
Mismatches:	132
Indels:	34
Gaps:	12

[illegible]

```

355 GAACCTGCACCTAATCTAGAGAAATCTCTAAATATAAACCGGACTTAAATGTGCGGTCA 414
128 AsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeu 147
415 AAAGTTAGAAATGAAAAAGTTTACGATCAATATCTAAATCGCACCAACAGTT----- 468
148 LysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAla 167
469 ---TCTACTGATACAGTTTCCAAATTCAAA---GATACAACTAAGTTAAATGGGAAAGCT 522
168 LeuGlyLysGluGluGluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGlu 187
523 TTAGGGAAGAAAAGAGAGCTGAAGATTACTTAAAGAGTACGATGATAAGTAGCTGCA 582
188 TyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLys 207
583 TTCAAAAA-----GATGCAAAAGCAAGTATAAGATGATGCCATTGAAA 630
208 SerGlyLeuLeuAlaHisProSerAsn-----SerTyrValGly 220
631 GCTTCAGTTGTTAACTTCCTGCTGATCATACAGAAATTTATGCTGGTGATGCTGGT 690
221 GlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeu 240
691 GAAATCTTAATGATTTAGGATTCAAACGT-----AATAAAGACTTA 732
241 SerLysTyrLeuLysGly-----ProTyrLeuGlnMetAsnThr---GluThrLeuSer 257
733 CAAAAACAAGTGTATGATGTTAAAGATATTTATCCAACTTACATCTAAGAAACATCCCA 792
258 GlnValAsnProGluArgMetPheIleMetThrAsnLysAlaSerSerAsnGluProSer 277
793 TTAATGAACCTGATCATATTTTGTAGTAAATCAGATCCAAATCGAAAGATGCTGCA 852
278 Leu-----LysGluLeuGluLysAspProValTyrLysLysLeuAsnAla 292
853 TTAGTAAAAAGACTGAAAGCGAAGTCACTTCAAGTAAAGAGTGGAAAAATTTAGACGCA 912
293 ValLysAsnGlnArgVal---AspIleLeuAspArgAspLeuTyrAlaArgSerArgGly 311
913 GTTAAAAACAACCAAGTATCTGATGATTTAGATGAAATCATTGGAACTTAGCTGCGCA 972
312 LeuIleSerSerGluGluMetAlaLysGluLeuValGluLeuSerLysLysAspSerLys 331
973 TATAAATCTTCAATAAACTTATTGACGATTTATATGAA-----AAGTTAATATTGAA 1026
332 LysAspAsnLys 335
1027 AAACAATCAAAA 1038

```

```

RESULT 24
US-10-329-624-238
; Sequence 238, Application US/10329624
; Publication No. US20040043037A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
;           Gil H. Choi
;           Patrick S. Dillon
;           Craig A. Rosen
;           Steven C. Barash
;           Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

```

```

; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/329,624
; FILING DATE: 27-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/956,171
; FILING DATE: October 20, 1997
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 238:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3775 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 238:
US-10-329-624-238

```

```

Alignment Scores:
Pred. No.: 1-67e-24 Length: 3775
Score: 405.00 Matches: 113
Percent Similarity: 51.74% Conservative: 65
Best Local Similarity: 32.85% Mismatches: 132
Query Match: 23.96% Indels: 34
DB: 19 Gaps: 12

```

US-10-724-972A-6352 (1-335) x US-10-329-624-238 (1-3775)

```

QY 8 LeuLysIleLeuSerValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCysGly 27
Db 61 ATTAATAATGCTT---GTGTAGCTTCTCTTCTACTTGT---TTACGAGATGAGT 114
QY 28 AsnAsnSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 47
Db 115 GGAATTTCAATAAACAATCATCTCATACAAAGATAAGGAAACAACATTTCAATTAACAT 174
QY 48 GluGluGlyThrThrLysValProLysHisProLysArgValValValValValValVal 67
Db 175 GCAATGGGTACAACTGAAATTAAGGGAAACCAAGCGTGTGTGTACGCTATATCAAGGT 234
QY 68 PheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLys 87
Db 235 GCCACTGACGTCGCTGATCTTTAGGTCTTAACTGTAGGTGCTGTAGATCATCGGACA 294
QY 88 LysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThr 107
Db 295 CAAAACCGAAATTCGAATACATATAAAATGATTTAAAGATATACTAAGATTTGAGTCAA 354
QY 108 ArgLysGlnProAsnLeuGluIleSerLysLysLysProAspLeuIleAlaAsp 127
Db 355 GAACCTGCACCTAATCTAGAGAAATCTCTAAATTAACACCGGACTTAAATTTGTCGCTCA 414
QY 128 AsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeu 147
Db 415 AAAGTTAGAAATGAAAGATTTAGATCAATATCTAAATCGCACCAACAGTT----- 468
QY 148 LysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAla 167
Db 469 ---TCTACTGATACAGTTTCCAAATTCAAA---GATACAACTAAGTTAAATGGGAAAGCT 522
QY 168 LeuGlyLysGluGluGluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGlu 187

```


QY 230 GluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeu 249
Db 691 -----TTAACAATACTTTTGAACGAGGAAATCTGAG-----CCTGATGGTTTCAA 738
QY 250 GlnMetAsnThrGluThrLeuSerGlnValAsnProGluArg---MetPheIleMetThr 268
Db 739 CAAACAACCTGGGAATCATTAACAAGGTACAAAGTCAAACTTCATTTACATTTGATGCG 798
QY 269 AsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGluLysAspProValTyrLys 288
Db 799 GATGAAGATAACATTTTGTGACACCACTAAATA-----GGCAACCTGCTTGGAA 849
QY 289 LysLeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTyrPalaArg 308
Db 850 GAATTAAGTTTAAAGAAGAAACAAATGTATATAATTAAGAAGCGGACACTTGGATTTC 909
QY 309 SerArgGlyLeuLysSerSerGluGluMetAlaLysGluLeuValGluLeu 325
Db 910 GGT---GGTCTGAGTCTGCAACATCTTTAGCAACACAGTAGCATGTA 957

RESULT 26

US-09-974-300-2061
; Sequence 2061, Application US/09974300
; Patent No. US20020146721A1

GENERAL INFORMATION:

APPLICANT: Berka, Randy M.

APPLICANT: Clausen, Ib Groth

TITLE OF INVENTION: Methods For Monitoring Multiple Gene

FILE REFERENCE: Expression

CURRENT APPLICATION NUMBER: 10085.500-US

PRIOR FILING DATE: 2001-10-05

PRIOR APPLICATION NUMBER: 09/680.598

PRIOR FILING DATE: 2000-10-06

PRIOR APPLICATION NUMBER: 60/279.526

PRIOR FILING DATE: 2001-03-27

NUMBER OF SEQ ID NOS: 8481

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2061

LENGTH: 944

TYPE: DNA

ORGANISM: Bacillus licheniformis

US-09-974-300-2061

Alignment Scores:

Pred. No.: 3.9e-23 Length: 944

Score: 380.50 Matches: 109

Percent Similarity: 52.92% Conservative: 63

Best Local Similarity: 33.54% Mismatches: 128

Query Match: 22.51% Indels: 25

DB: 9 Gaps: 13

US-10-724-972A-6352 (1-335) x US-09-974-300-2061 (1-944)

QY 8 LeuLysIleLeuSer---ValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCys 26
Db 1 TTGAAAAAATATCTTTATTGATTGATGCAATGATAACCGTTCTTGTGTCGAGCCGCTGC 60
QY 27 GlyAsnAsnSerSerAsnSerSerLysGluSerSer---LysAspGlyValGluIle 45
Db 61 GGAACAAGCTGCTGACACAGATGCAAAATCTTCTCTTAAAGAAAGAAACCGTCACAATT 120
QY 46 LysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeuGlu 65
Db 121 AAACAACAAGCGGACACACAGATGTACCGAAACCCGAAAGAAAGTGTGTCTTCAAT 180
QY 66 TyrSerPheValAspAlaLeu-----ValAlaLeuAspValLysProValGlyIleAla 83
Db 181 TTTGGAATGCTTGACACCGCTTGATGAGCTTGGCTTAAGCGACCGGTTGTGCGGCTTCG 240
QY 84 AspAspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThr 103
Db 241 AAACAACCCCTTCGCTCT---TACTTTGAAAAAATATAAAGAC-----AGCAAAATTTGAA 291

```

/ Publication No. US20050196754A1
/
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq, Inc
/ TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
/
/ FILE REFERENCE: 790CIP3/US
/ CURRENT APPLICATION NUMBER: US/10/450,763
/ PRIOR FILING DATE: 2003-06-11
/ PRIOR APPLICATION NUMBER: PCT/US01/08631
/ PRIOR FILING DATE: 2001-03-30
/ PRIOR APPLICATION NUMBER: 09/540,217
/ PRIOR FILING DATE: 2000-03-31
/ PRIOR APPLICATION NUMBER: 09/649,167
/ PRIOR FILING DATE: 2000-08-23
/ NUMBER OF SEQ ID NOS: 60736
/ SOFTWARE: Custom
/ SEQ ID NO 8842
/ LENGTH: 4392
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/
/ FEATURE:
/ NAME/KEY: SIMILAR
/ LOCATION: (1)..(654)
/ OTHER INFORMATION: 98% homologous to Escherichia coli K12 citrate-dependent iron
/ OTHER INFORMATION: transport, periplasmic protein, accession number AE000499, Smith-
/ OTHER INFORMATION: Waterman Score=1060.
/
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)..(4392)
/ OTHER INFORMATION: n = a,t,c or g
/
/ US-10-450-763-8842
/
/ Alignment Scores:
/ Pred. NO.: 7,58e-22 Length: 4392
/ Score: 375.00 Matches: 96
/ Percent Similarity: 50.17% Conservative: 52
/ Best Local Similarity: 32.54% Mismatches: 115
/ Query Match: 22.15% Indels: 32
/ DB: 24 Gaps: 5
/
/ US-10-724-972A-6352 (1-335) x US-10-450-763-8842 (1-4392)
/
/ QY 45 IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu 64
/ DB 70 GTTCAGGACGAACACGGCACGTTAAACACTCGAAAAAAGCCACACGCGATTGTGTGCTG 129
/
/ QY 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84
/ DB 130 GAACTCTGTTCCCGATCGCTGGCGCGCGTGCATCCGATCGGTATTGCCGAC 189
/
/ QY 85 AspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSer 104
/ DB 190 GATAACGATGCAAAACGCATCTGCCGGAAGTCGTCGCGCACCTGAAACCGTGCGAGTCC 249
/
/ QY 105 ValGlyThrArgLysGlnProAsnLeuGluGluIleSerLysLysLysProAspLeuIle 124
/ DB 250 GTCGGAACGCGCGCGCAGCGACGCTGGAAGCCATTGCCGCTCTGAAACCCAGACCTGATC 309
/
/ QY 125 IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144
/ DB 310 ATTGCCGACAGCAGTCGCCATGCGGGGTTTACATCCCTTCGAGCAATTCGCCCGGTA 369
/
/ QY 145 IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164
/ DB 370 CTGCTGCTTAAGTCCCGCAACGAACCTACGCTGAAATTTGCAATTCGCGCTATCATC 429
/
/ QY 165 SerLysAlaLeuGlyLysGluGluGlyLysLysArgLeuGluGluHisAspLysLys 184
/ DB 430 GGCGAATGTTGGTAAAGAACGAGAGATGCGAGCAGCTCTGGAAACAACATAAGAGAGG 489
/
/ QY 185 IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal 204
/ DB 490 ATGGCGCAGTGGGCGACCCAG-----CTTCCCAAGGACACGCGTGGCCTTTGGCACA 543

```

Alignment Scores:			
Pred. No.:	3,27e-22	Length:	2209
Score:	50.170	Matches:	96
Percent Similarity:	35.0%	Conservative:	52
Best Local Similarity:	32.54%	Mismatches:	115
Query Match:	22.19%	Indels:	32
DB:	24	Gaps:	5
US-10-724-972A-6352 (1-335) x US-10-450-763-7183 (1-2209)			
Qy	45	IleLysHisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeu	64
Db	2140	GTTCAGGACGAAACACGGCAGCTTAAACCTCGAAAAACCGCAACAGCTGTGGTGTG	2081
Qy	65	GluTyrSerPheValAspAlaLeuAspValLysProValGlyLeuAsp	84
Db	2080	GAATCTCTGCTCGCCGATGGCTGGCGCGGTGGAGCTCATCCGATCGGTATTCGCCAC	2021
Qy	85	AspAsnLysLysAsnArgIleLysProLeuArgAspLysIleGlyLysTyrThrSer	104
Db	2020	GATNACGATGCMAAACCGCATCTCTGCCGAAGTCGTCGCACCTGTGAACCGTGGCAGTCC	1961
Qy	105	ValGlyThrArgLysGlnProAsnLeuGluGluLysSerLysLysProAspLeuLeu	124
Db	1960	GTCGGAACGGCGCGCAGCGAGCTTGGAGCCATTTGCCGCTCTGAAACACGACCTGATC	1901
Qy	125	IleAlaAspAsnAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr	144
Db	1900	ATTCCGCACAGCAGCTCCCATGCGGGGTTCATCGCTTGACGAAATTCGGCGGTATCATC	1841
Qy	145	IleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle	164
Db	1840	CTGCTGTTAGTCCGCAACGAAACCTACGCTGMAAATTTGCAATCTCGGCTATCATC	1781
Qy	165	SerLysAlaLeuGlyLysGluGluGluGlyLysLysArgLeuGluHisAspLysLys	184
Db	1780	GSCGAATGTTGGGTAAAGACGAGAGATGCGACACGCTGTGAACACATAAAGAGG	1721
Qy	185	IleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuProAlaVal	204
Db	1720	ATGGCGCAGTGGCGCAGCAG-----CTTCCMAAGGACACGCTGGCTTGGCACA	1667
Qy	205	AlaAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlnPheLeuSer	224
Db	1666	TCACGCGAACGCAATTCACCTGCATATCTCAGGACCTGGACCGCAGCTGCTGGCC	1607
Qy	225	GlnLeuGlyPheLysGluAlaLeuSerAspAspValThrLysGlyLeuSerLysTyrLeu	244
Db	1606	TCCTCTGGGGCTG-----AACGTCCCGCTGGCATG	1577
Qy	245	LysGlyProTyrLeuGlnMetAsnThrGluThrLeuSerGlnValAsnProGluArgMet	264
Db	1576	GGGGTGGCTTCATCCGCTCTCGCTGGTCCGCTGGCGGCTTCATGTCACCGCATGT	1517
Qy	265	PheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuGluLysAsp	284
Db	1516	TCTATCCATGCAATTACGGTTACATCAACACACCTGCTC-----	1475
Qy	285	ProValTyrLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeu-----	301
Db	1474	-----TGACCGCGGTTCGTGTGATCGCTCGCGTTCGGGTTCGAAATGACC	1421
Qy	302	-----AspArgAspLeuTyrAlaArgSerArgGlyLeuIleSerSerGlu	316
Db	1420	GACGAACCGGTGAAGATCGGAACCTGGTTGGGTTCCGCACAGCAAGCTGACAAAGAA	1361
Qy	317	-----GluMetAlaLysGluLeuValGluLeuSerLys	327
Db	1360	TACGATCACATTAAACAGCTTAACGATCTGCTGAACTGCTGAAA	1316
RESULT 28			
US-10-450-763-8842			
; Sequence 8842, Application US/10450763			

205 AlaAlaLysSerGlyLeuLeuAlaHisProSerAnSerTyrValGlyGlnPheLeuSer 224
 544 TCACGCGNACAGCAATTCACCTGATACCTCAGGAGACCTGGACCGGCGAGCTGCGCC 603
 225 GlnLeuGlyPheLysGluAlaLeuSerAspValThrLysGlyLeuSerLysTyrLeu 244
 604 TCTCTGGGCTG-----AACGTCCTCCGCTGCGCATG 633
 245 LysGlyProTyrLeuGlnMetAsnThrLysLeuSerGlnValAsnProGluArgMet 264
 634 GCGGCTCGCTCCATCGCTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 693
 265 PheIleMetThrAsnLysAlaSerSerAsnGluProSerLeuLysGluLeuLysAsp 284
 694 TCTATCATCATGATACGTTACATCAACACACACCTGCTCTC----- 735
 285 ProValTyrLysLysLeuAsnAlaValLysAsnGlnArgValAspIleLeu----- 301
 736 -----TGGACGCGCGGCTTCTGATCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 789
 302 -----AspArgAspLeuTyrAlaArgSerArgGlyLeuLysSerSerGlu 316
 790 GACGAAGCGCGTGAAGATGCAAACTGGTTGCGGTTCCGACACAGCAAGCTGACGAAGA 849
 317 -----GluMetAlaLysGluLeuValGluLeuSerLys 327
 850 TAGCATCACATTAAGACGTTAAGATCTGCTGAACCTGCTGAAA 894

RESULT 29
 US-10-450-763-13139
 ; Sequence 13139, Application US/10450763
 ; Publication No. US20050196754A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
 ; FILE REFERENCE: 790CIP3/US
 ; CURRENT APPLICATION NUMBER: US/10/450,763
 ; CURRENT FILING DATE: 2003-06-11
 ; PRIOR APPLICATION NUMBER: PCT/US01/08631
 ; PRIOR FILING DATE: 2001-03-30
 ; PRIOR APPLICATION NUMBER: 09/540,217
 ; PRIOR FILING DATE: 2000-03-31
 ; PRIOR APPLICATION NUMBER: 09/649,167
 ; PRIOR FILING DATE: 2000-08-23
 ; NUMBER OF SEQ ID NOS: 60736
 ; SOFTWARE: Custom
 ; SEQ ID NO 13139
 ; LENGTH: 2799
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SIMILAR
 ; LOCATION: (181)..(903)
 ; OTHER INFORMATION: 73% homologous to Mus musculus Pro-Pol-dutPase
 ; OTHER INFORMATION: polyprotein, accession number Y12713.Smith-Waterman Score=892.
 US-10-450-763-13139

Alignment Scores:
 Pred. No.: 6,14e-20 Length: 2799
 Score: 350.00 Matches: 69
 Percent Similarity: 68.92% Conservative: 33
 Best Local Similarity: 46.62% Mismatches: 46
 Query Match: 20.71% Indels: 0
 DB: 24 Gaps: 0
 US-10-724-972A-6352 (1-335) x US-10-450-763-13139 (1-2799)
 Qy 45 IleLysHisGluGlyThrLysValProLysHisProLysArgValValValLeu 64
 Db 1672 GTTCAGGACGACACGCGGCTTTACATCTGAAAGAACCCACACGATGCTGCTG 1731
 Qy 65 GluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAsp 84

1732 GAACCTCTGTTGCGCGATGCGTGGCGCGCTGGACGTCATCCCGATCGGTATTGCGCGAC 1791
 Qy 85 AspAsnLysLysAsnArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSer 104
 1792 GATAACCATCAAAACGATCTCTCCCGAAGTGGTGGACCTCTGAAACCGTGGCAGTCC 1851
 Qy 105 ValGlyThrArgLysGlnProAsnLeuGluLysSerLysLysLysProAspIleLeu 124
 1852 GTCCGAACGCGCGCGACCGGCTGGAAGCCATTCGCGCTCTGAAACCGACGACCTGATC 1911
 Qy 125 IleAlaAspAsnArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThr 144
 1912 ATTCCGACACGACGTCGCTGCGGGGTTTACATCCCTTGGCAGCAATCGCGCGCGTA 1971
 Qy 145 IleGluLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLysThrIle 164
 1972 CTCTGCTTTAAGTCCCGACACGAAACCTACGCTGAAATTTGCAATCTGCGGCTATCATC 2031
 Qy 165 SerLysAlaLeuGlyLysGluGluGlyLysArgLeuGlyLysLysLysLysLysLys 184
 2032 GCGCAATGGTGGTAAAGCCGAGAGATGACGCGACGCTCTGGAACAAATAAGAGAGG 2091
 Qy 185 IleGluGluTyrLysLysGluIle 192
 2092 ATGGCGCAGTGGCGCCAGCAGCTT 2115

RESULT 30
 US-08-781-986A-604
 ; Sequence 604, Application US/08781986A
 ; Publication No. US20030054436A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 5255
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/781,986A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Benson, Bob
 ; REGISTRATION NUMBER: 30,446
 ; REFERENCE/DOCKET NUMBER: PB248PP
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 604:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2115 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; US-08-781-986A-604
 Alignment Scores:
 Pred. No.: 1.03e-18 Length: 2115
 Score: 334.00 Matches: 67
 Percent Similarity: 63.84% Conservative: 46
 Best Local Similarity: 37.85% Mismatches: 58

Query Match: 19.76% Indels: 6
DB: 8 Gaps: 3

US-10-724-972A-6352 (1-335) x US-08-781-986A-604 (1-2115)

QY 154 TyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 173
DB 3 TATAATGCAATATTGAAGCATTTAAACACAGTCCTAAAGCAGTAGGCAAGAGAA 62

QY 174 GlyLysLysArgLeuGluGluHisAspLysLysLysLysLysLysLysLysLys 193
DB 63 GGCAGAGAGCTCTGAAAGCATGATAAATATTAGCGGAGATTAGAAAGAAATTTGAA 122

QY 194 MetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLysLysLys 213
DB 123 CAGAGTACGTTAAATCTGCATTTGCATTCGGTATCTCAAGACGAGTATGTTTAAAT 182

QY 214 ProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSer 233
DB 183 AATGAAGATACATTTATGGGACAAATCTTAATTAATGAGTATTCACCTGAGTCAMA 242

QY 234 AspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThr 253
DB 243 AAARMAAACTACGCAATGTTGGTGAACGCAAGGCTGCTTATATATATATTTAAATAAT 302

QY 254 GluThrLeuSerGlnValAsnProGluArgMetPheLysMetThrAsn---LysAlaSer 272
DB 303 GAAGAACTTGCCAAATATCAATCCAAAGATTTATGATTTAGCCACTGACGGAACCGAC 362

QY 273 SerAsnGluProSerLeuLysGluLeuGlyLysAspPro---ValTyrLysLysLys 291
DB 363 AAAAATAGAACGAAATTC-----ATTGATCTCGAGTTTGGAAATCATTTAAAA 410

QY 292 AlaValLysAsnGlnArgValAspIleLeuAspArgLeuTyrAlaArgSerArgGly 311
DB 411 GCTGTGAAGATAACAAAGTTTATGACGTTGACCGAATAGTGTTCGAATCAAGGGG 470

QY 312 LeuLysSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328
DB 471 ATTATCGCAAGTGAAGTATGCGAGAAGATTTAGAAAAAATTTGCAGAAAAA 521

RESULT 31

US-10-329-624-604
; Sequence 604, Application US/10329624
; Publication No. US20040043037A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Rannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/329,624
; FILING DATE: 27-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/956,171
; FILING DATE: October 20, 1997
; APPLICATION NUMBER: 60/009,861

FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 604:
SEQUENCE CHARACTERISTICS:
LENGTH: 2115 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 604:
US-10-329-624-604

Alignment Scores: Length: 2115
Pred. No.: 334.00 Matches: 67
Score: 63.84% Conservative: 46
Percent Similarity: 37.85% Mismatches: 58
Best Local Similarity: 19.76% Indels: 6
Query Match: 19 Gaps: 3
DB: 3

US-10-724-972A-6352 (1-335) x US-10-329-624-604 (1-2115)

QY 154 TyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGluGlu 173
DB 3 TATAATGCAATATTGAAGCATTTAAACACAGTCCTAAAGCAGTAGGCAAGAGAA 62

QY 174 GlyLysLysArgLeuGluGluHisAspLysLysLysLysLysLysLysLysLys 193
DB 63 GGCAGAGAGCTCTGAAAGCATGATAAATATTAGCGGAGATTAGAAAGAAATTTGAA 122

QY 194 MetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLysLysLys 213
DB 123 CAGAGTACGTTAAATCTGCATTTGCATTCGGTATCTCAAGACGAGTATGTTTAAAT 182

QY 214 ProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAlaLeuSer 233
DB 183 AATGAAGATACATTTATGGGACAAATCTTAATTAATGAGTATTCACCTGAGTCAMA 242

QY 234 AspAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMetAsnThr 253
DB 243 AAARMAAACTACGCAATGTTGGTGAACGCAAGGCTGCTTATATATATATTTAAATAAT 302

QY 254 GluThrLeuSerGlnValAsnProGluArgMetPheLysMetThrAsn---LysAlaSer 272
DB 303 GAAGAACTTGCCAAATATCAATCCAAAGATTTATGATTTAGCCACTGACGGAACCGAC 362

QY 273 SerAsnGluProSerLeuLysGluLeuGlyLysAspPro---ValTyrLysLysLys 291
DB 363 AAAAATAGAACGAAATTC-----ATTGATCTCGAGTTTGGAAATCATTTAAAA 410

QY 292 AlaValLysAsnGlnArgValAspIleLeuAspArgLeuTyrAlaArgSerArgGly 311
DB 411 GCTGTGAAGATAACAAAGTTTATGACGTTGACCGAATAGTGTTCGAATCAAGGGG 470

QY 312 LeuLysSerGluGluMetAlaLysGluLeuValGluLeuSerLysLys 328
DB 471 ATTATCGCAAGTGAAGTATGCGAGAAGATTTAGAAAAAATTTGCAGAAAAA 521

RESULT 32

US-08-781-986A-2556
; Sequence 2556, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255

;;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 2556:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 242 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-781-986A-2556

Alignment Scores:
Pred. No.: 8,86e-20 Length: 242
Score: 333.00 Matches: 66
Percent Similarity: 90.00% Conservative: 6
Best Local Similarity: 82.50% Mismatches: 8
Query Match: 19.70% Indels: 0
DB: Gaps: 0

US-10-724-972A-6352 (1-335) x US-08-781-986A-2556 (1-242)
QY 70 AspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLysLysAen 89
Db 2 GATGATTAGCAGCATTAGACGTTAAACCCAGTTGGTATTGCTGATGATGTAAGAAAAA 61
QY 90 ArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLys 109
Db 62 CGTATCATTAACCCAGTTAGAGAAAAATTGGGATTATATCTCTGATGATACGTA 121
QY 110 GlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsnAen 129
Db 122 CAGCCAAACTTAGAAGAAATTAGTAAATTAACCCGATTATTAATTCGCTGATACG 181
QY 130 ArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSer 149
Db 182 AGACATANAGGTATTAAATAAGAAATTAAACAAATTTGCCAACCATTTATCATTAAGAAGT 241
RESULT 33
US-10-329-624-2556
; Sequence 2556, Application US/10329624
; Publication No. US2004004037A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256

;;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/329,624
; FILING DATE: 27-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/956,171
; FILING DATE: October 20, 1997
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248PID1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 2556:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 242 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-10-329-624-2556

Alignment Scores:
Pred. No.: 8,86e-20 Length: 242
Score: 333.00 Matches: 66
Percent Similarity: 90.00% Conservative: 6
Best Local Similarity: 82.50% Mismatches: 8
Query Match: 19.70% Indels: 0
DB: Gaps: 0

US-10-724-972A-6352 (1-335) x US-10-329-624-2556 (1-242)
QY 70 AspAlaLeuValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLysLysAen 89
Db 2 GATGATTAGCAGCATTAGACGTTAAACCCAGTTGGTATTGCTGATGATGTAAGAAAAA 61
QY 90 ArgIleIleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLys 109
Db 62 CGTATCATTAACCCAGTTAGAGAAAAATTGGGATTATATCTCTGATGATACGTA 121
QY 110 GlnProAsnLeuGluGluIleSerLysLeuLysProAspLeuIleAlaAspAsnAen 129
Db 122 CAGCCAAACTTAGAAGAAATTAGTAAATTAACCCGATTATTAATTCGCTGATACG 181
QY 130 ArgHisLysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSer 149
Db 182 AGACATANAGGTATTAAATAAGAAATTAAACAAATTTGCCAACCATTTATCATTAAGAAGT 241
RESULT 34
US-09-738-626-375
; Sequence 375, Application US/09738626
; Publication No. US2002019760S1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO

QY 212 AlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLeuGlyPheLysGluAla 231
Db 634 TATGGTCCC---GAGACCTTCTCTGTTTCTGAGTCTTCTGAAACAAGTTGGATTT-----681
QY 232 LeuSerAspValThrLysGlyLeuSerLysValTyrLeuLysGlyProTyrLeuGlnMet 251
Db 682 -----GACCTGGGGAACGTGATTGGAATGAGTAC-----TCCATGATGGAGCTA 726
QY 252 AsnThrGluThrLeuSerGlnValAsnProGluArgMetPhe-----IleMetThrAsn 269
Db 727 TCTCTGAAAACCTTTGGCGAGATCGATGGAGACCTTATTCTTACACCATCCAGGATCC 786
QY 270 LysAlaSerSerAsnGluProSerLeuLysGluLeuLysAspProValTyrLysLys 289
Db 787 CCTGAAGCAACCACTTATCCAAAGATTTCCGAACTG-----TGGTTGAT 831
QY 290 LeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTyrAlaArgSer 309
Db 832 TCACCGACAGTTCGGCAAGGTAGAACTTACGAGTTTGAACGACCAACCTGGATCGTCGC 891
QY 310 ArgGlyLeuIleSerSerGluGluMetAlaLysGluLeuValGlu 324
Db 892 ATCGGTGTATTAGTGCCTAATGAAATCTTGGATGACCTCGAAGAA 936
RESULT 35
US-09-738-626-1
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR FILING DATE: 1999-12-16
; PRIOR FILING DATE: 2000-04-07
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1
Alignment Scores:
Pred. No.: 1,518-14 Length: 3309400
Score: 331.00 Matches: 90
Percent Similarity: 49.21% Conservative: 65
Best Local Similarity: 28.57% Mismatches: 140
Query Match: 19.59% Indels: 20
DB: Gaps: 7
US-10-724-972A-6352 (1-335) x US-09-738-626-1 (1-3309400)
QY 13 ValIleGlyLeuPheValLeuIleAlaThrAlaAlaCysGlyAsnAsnSerSer 32
Db 352735 GTGGCCCGCTCTCTCGGCGTCGCACCTCTTAAAGCTCTCTTCAACATCTTCCGAT 352794
QY 33 AsnSer---SerLysGluSerSerLysAspGlyValGluLysHisGluGlyThr 51

APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR FILING DATE: 1999-12-16
PRIOR FILING DATE: 2000-04-07
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 375
LENGTH: 945
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
US-09-738-626-375
Alignment Scores:
Pred. No.: 6,978-19 Length: 945
Score: 331.00 Matches: 90
Percent Similarity: 49.21% Conservative: 65
Best Local Similarity: 28.57% Mismatches: 140
Query Match: 19.59% Indels: 20
DB: Gaps: 7
US-10-724-972A-6352 (1-335) x US-09-738-626-375 (1-945)
QY 13 ValIleGlyLeuPheValLeuIleAlaThrAlaAlaCysGlyAsnAsnSerSer 32
Db 43 GTGGCCCGCTCTCTCGGCGTCGCACCTCTTAAAGCTCTCTTCAACATCTTCCGAT 102
QY 33 AsnSer---SerLysGluSerSerLysAspGlyValGluLysHisGluGlyThr 51
Db 103 GAATCAATCAACCTGAGTTCGCCAGCATCGATATTCTGAGCAGCAGCAATGGCGACC 162
QY 52 ThrLysValProLysHisProLysArgValValLeuGluTyrSerPheValAspAla 71
Db 163 ACCGAAATCCCTGAAACCCACCGGTGGTGGTCGTCATGATCCCCACACCTCCAGCGCA 222
QY 72 LeuValAlaLeuAspValLysProValGlyIleAlaAspAspAsnLysLysArgIle 91
Db 223 CTTTGGCTTTGGGAATTAATCTCAGTCGAGCTACGGAATCTGGATCCGAAATGTTTC 282
QY 92 IleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnPro 111
Db 283 CCGCGCTACTTGGCTGACGAGCTAAAGACACCGAATCTGTTGGGCTGACATCTGAGCCA 342
QY 112 AsnLeuGluGluIleSerLysLysLysProAspLeuIleAlaAspAsnAsnArgHis 131
Db 343 AATTGGAAAGATCCCGACCTGGATCCGGATTTGATCATTTGGCCCAAGGTCCGCCAC 402
QY 132 LysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAsp 151
Db 403 GAGGCTATTATGATCAGCTTTCCAGACATCGCACCACCGTGTGTCGAAAGTTCCGCGC 462
QY 152 GlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuGlyLysGlu 171
Db 463 ACAAACTGGAATGAACAGGAGAA-----ATCACTCGCGCGACGAGTAAACAAGTCT 513
QY 172 GluGluGlyLysLysArgLeuGluGluHisAspLysLysIleGluGluTyrLysGlu 191
Db 514 GATGAGATGACAAATCTGATCTCAGACTTGGACACCGGTCGCCAGAGCTTGTGTGAGAG 573
QY 192 IleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeu 211
Db 574 ATCGGTGCTGACGGACAAACCGCTTCAATGGTTTCCGATTCGCCGACCAACTTCCAGGCTC 633

Db 352795 GAATCAATCCAACTGAAAGTTGCCAGCACTGGATATTAGTGGAGCACCAATGGGCACC 352854
Qy ThrLysValProLysHisProLysArgValValValValLysGluTyrSerPheValAspAla 71
Db 352855 ACCGAAATCCCTGAAACCCCAACCGGTCTGCTCTGATGATTCCTCCACACCTCGACGCA 352914
Qy 72 LeuValAlaLeuAspValLysProValGlyLeuAlaLeuAspAspAspAspAspAspAsp 91
Db 352915 CTTTGGCTTTGGGAAATTACTCCAGTCGGAGCTACGGAATCTGATCCGAAATGGTTTC 352974
Qy 92 IleLysProLeuArgAspLysIleGlyLysTyrThrSerValGlyThrArgLysGlnPro 111
Db 352975 CCCGCTACTTGGCTGACGAGCTAAAGACACCAACCGATCTGTGGCTCGACATCTGACCCA 353034
Qy 112 AsnLeuGluGluLysSerLysLeuLysProAspLeuIleAlaLeuAspAsnAsnAtgHis 131
Db 353035 AATTGGAAAGATCGCGCACTGGATCCGATTCGATTCGCGCAAGGTCGCCAC 353094
Qy 132 LysGlyIleTyrLysAspLeuAsnLysIleAlaProThrIleGluLeuLysSerPheAsp 151
Db 353095 GAGGCTATTATGATCAGCTTTCAGACATCGCAACCGTGATGTCGGAAGGTTCCGGC 353154
Qy 152 GlyAspTyrAsnGluAsnIleAspAlaPheLysThrIleSerLysAlaLeuLysGlu 171
Db 353155 ACAACTGGATGAACAGCGAGAA-----ATCACTCGCGCAGCAGTAAACAAGTCT 353205
Qy 172 GluGluGlyLysLysArgLeuGluGluHisAspLysIleGluGluTyrLysLysGlu 191
Db 353206 GATGAGATGACAACTGATCTCAGACTTGGACACCGTCCACAGCTTGTGTGAAGAG 353265
Qy 192 IleThrMetAspLysAsnGlnLysValLeuProAlaValAlaLysSerGlyLeuLeu 211
Db 353266 ATCGGTGCTGACGCAACCGCTTCAATGTTGATTCGATTCGCGACGCAACTTCAGGCTC 353325
Qy 212 AlaHisProSerAsnSerTyrValGlyGlnPheLeuSerGlnLysGlyPheLysGluAla 231
Db 353326 TATGTTCC---GAGACCTTCTCTGTTTCTGATGTTCTGGAACAAGTTGGATT----- 353373
Qy 232 LeuSerAspValThrLysGlyLeuSerLysTyrLeuLysGlyProTyrLeuGlnMet 251
Db 353374 -----GACTCGGGGAACGTGATGGAATGATGAC-----TCCATGATGGAGCTA 353418
Qy 252 AsnThrGluThrLeuSerGlnValAsnProGluArgMetPhe-----IleMetThrAsn 269
Db 353419 TCTCAGAAACTTTGGGAGATCGATGAGACCTTATTTTCTACACCATCCAGGATCC 353478
Qy 270 LysAlaSerSerAsnGluProSerLysGlyLeuGluLysAspProValTrpLysLys 289
Db 353479 CCTGAAGCAACCACTTATCCAAAGATTTCCGAAGT-----TGGGTTGAT 353523
Qy 290 LeuAsnAlaValLysAsnGlnArgValAspIleLeuAspArgAspLeuTrpAlaArgSer 309
Db 353524 TCACAGCAGTTCCGCAAGTGAAGTACCTAGGATTTGAAGCAACCTGGATGTCGCGC 353583
Qy 310 ArgGlyLeuLysSerGluGluMetAlaLysGluLeuValGlu 324
Db 353584 ATCGGTGATTAGTGCCCAATGAAATCTTGGATGACCTTGGAGAA 353628

RESULT 36

US-09-974-300-2087

; Sequence 2087, Application US/09974300

; Patent No. US20020146721A1

; GENERAL INFORMATION:

; APPLICANT: Berka, Randy M.

; APPLICANT: Clausen, Ib Groth

; TITLE OF INVENTION: Methods For Monitoring Multiple Gene

; FILE REFERENCE: Expression

; CURRENT APPLICATION NUMBER: US/09/974,300

; PRIOR FILING DATE: 2001-10-05

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2087
; LENGTH: 648
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-2087

Alignment Scores:

Pred. No.: 6,52e-19 Length: 648
Score: 329.00 Matches: 74
Percent Similarity: 55.34% Conservative: 40
Best Local Similarity: 35.92% Mismatches: 88
Query Match: 19.47% Indels: 4
DB: 9 Gaps: 2

US-10-724-972A-6352 (1-335) x US-09-974-300-2087 (1-648)

Qy 24 AlaAlaCysGlyAsnAsnSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 42
Db 27 GCGGCTGCGGAAATCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCT 86
Qy 43 ValGluIleLysHisGluGlyThrThrLysValProLysHisProLysHisProLysHisPro 62
Db 87 TATACGTGAAGCACCAATGGCAACGCGCAATTTGACGGCACTCCGAAAAAAGTCGTC 146
Qy 63 ValLeuGluTyrSerPheValAspAlaLeuValAlaLeuAspValLysProValGlyIle 82
Db 147 GTACTCAGCAATGAAGCAGACAGACAGCTGCTCGCACTGGCGTAAACCTGTCGGCGCC 206
Qy 83 AlaAspAspAsnLysLysAsnArgIleLysProLeuArgAspLysLysLysLysLysLys 102
Db 207 GTTCAATCTCGCTGGAGATCATGTTATGACCATATTAAAGATAAAATGAAAGCGCTA 266
Qy 103 ThrSerValGlyThrArgLysGlnProAsnLeuGluGluLysSerLysLysLysLysLys 122
Db 267 GAGACGTGCGGAGAGGCTGCGCGCAATATTGAAAAAATCGGAGCCTGAAGCCTGAC 326
Qy 123 LeuIleAlaAspAsnAsnArgHisLysGlyLysLysLysLysLysLysLysLysLysLys 142
Db 327 TTAATCATCGCAACAAAATCGCAAGAGAGTATACGCAAGCTGAAAGCAATCGCG 386
Qy 143 ProThrIleGluLeuLysSerPheAspGlyAspTyrAsnGluAsnIleAspAlaPheLys 162
Db 387 CCGACGCTATTGCGGAGACATTTGTCAGGGGAATCGAAAAAACAAC-----TTTAAG 437
Qy 163 ThrIleSerLysAlaLeuGlyLysGluGluGlyLysLysLysLysLysLysLysLysLys 182
Db 438 CTTACTCGGAGCGCTCAATAAAAAAGCTGAAGCGGAAAGTCAATCGCTGACTTTGAC 497
Qy 183 LysLysIleGluGluTyrLysLysGluIleThrMetAspLysAsnGlnLysValLeuPro 202
Db 498 AAGCGCTGTAAGACATTCATGAAAAACTAGCGGTAAGTAAAGTAAACAAAAAGTCTCTCTC 557
Qy 203 AlaValAlaLysSerGlyLeuLeuAlaHisProSerAsnSerTyrValGlyGlnPhe 222
Db 558 GTCCGCTTTACAGATGCTGATACCGCATCTATCACAAGGATCATTTGCGGGAACGATT 617
Qy 223 LeuSerGlnLeuGlyPhe 228
Db 618 TTGGATCAGCTCGGTTTT 635

RESULT 37

US-10-282-122A-17083

; Sequence 17083, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

Qy	256	LeuSerGlnValAsnProGluArgMetPhe-----lleMetThrAsnLysAlaSerSer	273
Db	686	TTTGGGAGATCGATGGAGACCTTTATTTCACACCATCCAGGATCCCTCAAGCAACC	745
Qy	274	AsnGluProSerLeuLysGluLeuGluLysAspProValTrrpLysLysLeuAsnAlaVal	293
Db	746	ACTTATCCAAGATVTTCCGAAGT-----TGGGTGTGATTCACGACGAGTT	790
Qy	294	LysAsnGlnArgValAspIleLeuAspArgAspLeuTrrpAlaArgSerArgGlyLeuIle	313
Db	791	CGGCAAGGTAGAACTTACGAGTTTGAGAGCAAACTGGATGGTCCGGCATCGGTGTATT	850
Qy	314	SerSerGluGluMetAlaLysGluLeuValGlu	324
Db	851	GGTGCCCAATGAAATCTTGGATGACCTGGAAGAA	883

APPLICANT: Clausen, Ib Groth
TITLE OF INFORMATION:

```

; ; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06

```

```

; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2044
; LENGTH: 957
;

```

life: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-2044

Alignment Scores:
Pred. No.: 5, 12e-18 Length: 957
Score:

Percent Similarity:	50.4%	Conservative:	68
Best Local Similarity:	29.7%	Mismatches:	131
Query Match:	18.9%	Indels:	32
DB:	9	Gaps:	13

QY: 9 LysIleLeuSerValIleGlyLeuLeuPheValLeuIleAlaThrAlaAlaCysGlyAsn 28
||||| :::: |||| :::: ||||
dbd 7 AAAATATCAATATTCTATTTATTCTCTTCGCGCTTGGCGACGGGGATCGCGGAAC 66

QY 29 Asn-----SerSerSerAsnSerSerLysGluSerSerLysAspGlyValGluIleLys 46

dbb 67 AGCAAGAAAAGGCTGGCGGGCATGGCAAGAGCACAGCGTCAGCTCAGAGAGAGCATCGAG 126
::: ::: ::::: ||| ::::: |||::: |||:::
2y 47 HisGluGluGlyThrThrLysValProLysHisProLysArgValValValLeuGlu--- 65
::: ::: ||||| ||||| ||||| ::|||::: |||:::
b 127 TACCTCGATCATATACACAAGGTG---AAGATCCCGGGCGGAACAATCGTCAATACGGGA 183

66 -----TyrSerPheValAspAlaLeuValAlaLeuAspValValProValGly---ile 82
184 AGCGTGAATCGATGGAAGATGCCAAGCTT---CTGGATGTTCCACCGCGGTGCCAAT 240
83 AlaAspAspAsnLysLysAsnArgileileLysProLeuArgAspLysileGlyLysTyr 102

241 TCCTTCTCCGAAAGTTCCCGGATCTGTGTTTAAGAGCATTA CTGACAAAGCA----- 291

103 ThrSerValGlyThrArgLysGlnProAsnLeuGluIleSerLysLeuLysProAsp 122
|||:::||||| :::: |||||||||:::||| ::||| |||||

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☐ BLACK BORDERS

☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES

☐ FADED TEXT OR DRAWING

☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING

☐ SKEWED/SLANTED IMAGES

☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS

☐ GRAY SCALE DOCUMENTS

☐ LINES OR MARKS ON ORIGINAL DOCUMENT

☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

This Page Blank (uspto)